AD-A007 585

IMAGE DESCRIPTORS FOR DISPLAYS

Roger W. Cohen, et al

RCA Laboratories

Prepared for:

Office of Naval Research

March 1975

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

AD-A007 585

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER	2. GOVT ACCESSION NO. / 3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle)	5. TYPE OF REPORT & PERIOD COVERED
IMAGE DESCRIPTORS FOR DISPLAYS	Technical Report (12-10-73 to 12-9-74)
	6. PERFORMING ORG. REPORT NUMBER: PRRL-75-CR-2
7. AUTHOR(s)	8. CONTRACT OR GRANT NUMBER(s)
Roger W. Cohen, Istvan Gorog, and Curtis R. Carlson	N00014-74-C-0184
9. PERFORMING ORGANIZATION NAME AND ADDRESS	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
RCA Laboratories Princeton, New Jersey 08540	Task No. NR-213-120
11. CONTROLLING OFFICE NAME AND ADDRESS	12. REPORT DATE
Office of Naval Research Arlington, Virginia 22217	March 1975
	13. NUMBER OF PAGES: 154
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)	15. SECURITY CLASS. (of this report): Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE: N/A

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Modulation Transfer Function (MTF)	Information Theory
Human Visual System	Information Capacity
Displays	Signal-to-Noise Ratio
Sampling	Perception

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This report summarizes the results of our experimental and theoretical efforts aimed at developing new image quality descriptors suitable for performance prediction and optimization of combined display-observer systems. Our approach is that of linear systems analysis, using the methodology of statistical communications theory. We treat the processes of image display and perception in statistical terms and

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

19.

Statistical Communication Theory
Power Spectrum
Statistical Ensembles
Contrast Sensitivity Function
Image Quality Descriptors
Optical Processing
Visual Capacity

20.

predict the most probable performance of the overall observer-imaging system. This report contains a number of important new results concerning the statistical properties of natural scenes, a complete, new, statistical treatment of the one-dimensional image sampling problem, a new descriptor suitable for the analysis of the perceived sharpness of two-dimensional analog (nonsampled) displays having nonisotropic point-spread functions, and data obtained from a set of psychophysical measurements aimed at the quantitative understanding of the two-dimensional sine wave response of human visual perception. We also report the development of a real-time optical block image processor that was developed for use for the experimental testing of our theoretical results.

PREFACE

This Technical Report was prepared by RCA Laboratories, Princeton, New Jersey 08540, for the Office of Naval REsearch under Contract No. N00014-74-C-0184, task number NR-213-120. It describes work performed during the period 10 December 1973 to 9 December 1974 in the Physical Electronics Research Laboratory, G. D. Cody, Director. The Project Scientist is R. W. Cohen.

The authors thank Dr. G. D. Cody for his continuous support, encouragement, and his useful suggestions; Dr. J. K. Clemens for his help, advice, and direct participation in the power spectral density measurements; Dr. A. Rose and Mr. J. J. Mezrich for their participation in the design and methodology of the psychophysical measurement portions of this program; Dr. P. M. Heyman and Mr. W. M. Moles for their advice during the electronic system design phases of the program; Mr. J. R. Woolston for his participation in the computer interfacing and software development phases of the program; Dr. P. A. Levine for making his CCD camera system available for our spectral analysis; Dr. D. O. North for frequent consultations about various aspects of linear systems analysis; and Dr. O. H. Schade for his advice concerning the use of TV monitors in perceptual measurements.

STANDARD OBSERVER RESPONSE

In the report that follows we show that the quantitative analysis of the performance of the display/observer system based on the display MTF and on the observer's sine wave response provides useful and easily applicable measures for performance prediction. However, in order to assure the general applicability of such measures, it is imperative that a set of "standard observer sine wave response curves" be established. A set, rather than a single curve, is required in order to take into account the strong dependence on average brightness of the sine wave response curve. We propose that under the auspices of the appropriate government agencies and professional societies an international committee be set up to undertake the specification of the standard observer sine wave response in a manner similar to that employed to establish the CIE standard observer color response. Such a committee would be charged with ensuring that proper measurement techniques and experimental precautions are employed, as well as the dissemination of the results to the relevant scientific, medical, and engineering institutions.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS (Continued)

LIST OF ILLUSTRATIONS (Continued)

LIST OF ILLUSTRATIONS (Continued)

LIST OF ILLUSTRATIONS (Continued)

SYMBOLS EMPLOYED IN FORMULAS AND TEXT

a	Dimensionless anisotropy factor [Eq. (126)]
A	Area of display
b	Grating contrast [Eq. (137)]
B	Power amplitude factor for the power spectrum $\Phi(\omega) = B/\omega^2$
C	Power amplitude factor for the white power spectrum $\Phi(\omega) = C$
C_v^T	Visual capacity (one-dimensional) [Eqs. (28), (29)]
C_{v2}^T	Two-dimensional visual capacity [Eq. (117)]
D	Strength of delta-function input (two-dimensional) [Eq. (111)]
$E_o(x)$	Perceived intensity pattern produced by a perfect display [Eq. (80)]
$E_m(x)$	Contribution to perceived intensity pattern from index value m [Eqs. (13),(14)]
$E_S(x)$	Perceived signal pattern [Eq. (16)]
$E_\delta(\rho,\alpha)$	Perceived two-dimensional intensity pattern due to point delta-function input [Eq. (113)]
$E(x)$	Perceived intensity pattern [Eq. (12)]
$E_N(x)$	Perceived noise pattern [Eq. (17)]
$\&$	Equivalence factor between sampled and analog displays with the same maximum H(r) [Eq. (108)]
F_s	Summation defined in Eq. (65)
G	Perceived gradient content [Eq. (36)]
h	Picture height for 4:3 aspect ratio television display
H	Total information capacity [Eq. (93)]
$I_F(x)$	Filtered input scene [Eq. (1)]
$I_o(x)$	Input scene
$\hat{I}_o(\omega)$	Spatial Fourier transform of input scene
I_n	Output of n'th sampling gate [Eq. (4)]
$I_\delta(\rho,\alpha)$	Two-dimensional intensity pattern on display screen due to point delta-function input [Eq. (111)]
$I(x)$	Intensity pattern on display screen [Eq. (5)]

K	Kell factor
L_o	Average luminance of displayed gratings
L_{max}	Maximum luminance of displayed gratings [Fig. 35]
L_{min}	Minimum luminance of displayed gratings [Fig. 35]
m	Index denoting contribution from m'th multiple of sampling frequency
n	Index denoting n'th sampling gate
N_{rms}	Root mean square number of perceived edges [Eq. (44)]
N_s	Number of samples across display: $N_s = \omega_s w/2\pi$ [Eq. (2)]
N_{TV}^{eq}	Number of TV lines for noiseless display needed to give same maximum H(r) as extremely noisy display [Eq. (107)]
N_{TV}	Limiting resolution expressed in TV lines: $N_{TV} = \omega_M w/\pi$
$N(\omega)$	Noise power spectrum
N_s^2	Average noise power (brightness squared) as measured on display screen
N_T^2	Total display-observer noise power (brightness squared) [Eq. (90)]
N_v^2	Internal noise power of human visual system (brightness squared)
N^2	Average perceived noise power (brightness squared) [Eq. (51)]
$O(\nu)$	MTF of human visual system (one-dimensional) [Fig. 1]
$O(\nu,\phi)$	Two-dimensional MTF of human visual system
p_c	Critical value of printing width [Eq. (78)]
p	Effective width of printing function, expressed in units of x_o
$P(x)$	Display printing function [Eq. (5)]
$\hat{P}(\omega)$	Spatial Fourier transform of printing function
q	Number of permissible pulse levels for a communication system
Q	Correlation quality [Eq. (79)]
r_{opt}	Viewing distance for maximum H
r_p	Viewing distance for maximum C_v^T
r	Viewing distance
$R_{eff}(\omega)$	Effective MTF of sampled display system [Eq. (23)]

$R_o(\omega)$	Overall MTF of analog display system (one-dimensional)
$R_o(\omega,\phi)$	Two-dimensional overall MTF of analog display system
$R(\omega)$	Initial band-limiting MTF of sampled display system [Eq. (1)]
s	Dimensionless sampling width [Eq. (4)]
S	Number of discrete states available to a communication system
S_o^2	Average perceived signal power (brightness squared) for perfect display [Eq. (85)]
S_s^2	Average signal power (brightness squared) as measured on display screen
S^2	Average perceived signal power (brightness squared) [Eq. (50)]
$(S/N)_h$	Effective signal-to-noise ratio of human visual system
$(S/N)_T$	Total system signal-to-noise ratio (Eq. (92)]
(S/N)	Perceived signal-to-noise ratio [Eqs. (50), (51)]
T	Message duration time
w	Display width
x_e	Perceived edge width [Eq. (25)]
x_o	Width of sampling gate [Eq. (3)]
x'	Distance coordinate along sine wave grating [Fig. 34]
x,y	Rectangular coordinate axes on display screen
α	Polar angle on display screen, as measured from the horizontal
δC_{v2}^T	Change in C_{v2}^T due to anisotropy [Eq. (136)]
δr_p	Change in viewing distance for maximum C_{v2}^T due to anisotropy [Eq. (135)]
δx	An arbitrary displacement of the input scene
ε	Mean square perceived error [Eq. (82)]
θ_e^∞	Perceived angular width of a single edge transition [Eq. (29)]
μ	Maximum pulse transmission rate of a communication system
ν_c	Effective cutoff frequency of human visual system
ν_{eff}	Effective retinal frequency for complex gratings
ν_o	Frequency for maximum $O(\nu)$: ν_o = 8 cycles/degree-of-vision [Fig. 5]
ν_x,ν_y	Retinal frequency coordinate along x and y axes, respectively
ν	Frequency coordinate on the retina: $\nu = \omega r/2\pi$

$\hat{\Pi}(\omega)$	Renormalized Fourier transform of printing function [Eq. (11)]
ϕ	Polar angle in two-dimensional frequency space, measured from the horizontal
$\Phi(\omega)$	Signal power spectrum for input scenes
ρ	Radial distance coordinate on screen
σ	Perceived spot area [Eq. (114)]
ω_L	Lower cutoff frequency of signal power spectrum
ω_{Mh}	Maximum display frequency in the horizontal direction
ω_{Mo}	Geometric mean of ω_{Mh} and ω_{Mv} [Eq. (125)]
ω_{Mv}	Maximum display frequency in the vertical direction
$\omega_M(\phi)$	Maximum frequency for which MTF $R_o(\omega,\phi)$ is non-zero [Eq. (128)]
ω_M	Maximum frequency for which MTF $R(\omega)$ or $R_o(\omega)$ is non-zero
ω_s	Sampling frequency: $\omega_s = 2\pi/x_o$ [Eq. (3)]
ω	Spatial frequency coordinate measured on display screen
Ω	Bandwidth of communication system
$\langle\ \rangle$	Denoting the operation of taking the average over an ensemble of scenes
$\langle\ \rangle^{-}$	Denoting the operation of taking the average over an ensemble of scenes and the average over the display screen.

SECTION I

INTRODUCTION

The ultimate goal of our image analysis program is to develop a formalism that can be employed to optimize the performance of sampled imaging systems, subject to constraints that may include both engineering and economic considerations. To achieve this goal we need to understand image sampling, to develop two-dimensional image quality descriptors, to obtain a detailed description of the performance of human visual perception, to combine our understanding of image sampling and visual perception into a unified theory suitable for performance prediction, and to test our theoretical predictions through suitable simulation experiments. This report contains our findings during the first phase of our image analysis program. It contains a number of new results concerning the statistical properties of natural scenes, a complete, new, statistical treatment of the one-dimensional image sampling problem, a new descriptor suitable for the analysis of the perceived shapness of two-dimensional analog (nonsampled) displays having nonisotropic point-spread functions, and data obtained from a set of psychophysical measurements aimed at the quantitative understanding of the two-dimensional sine wave response of human visual perception.

Our approach has been that of linear systems analysis, using the methodology of statistical communications theory. We are fully aware of the limitations of linear system analysis when it is used to study highly nonlinear systems such as human perception. Nevertheless, we know that all nonlinear physical systems can be described in a perturbation sense by linear equations, and thus far we found that the most important predictions obtained from our linear analysis can be verified experimentally. We treat the image sampling and sampled image perception in statistical terms. This is a significant departure from the frequently employed standard approach of calculating the wave shape of sampled simple test patterns. The standard approach allows one to gain an insight into the detailed mechanisms of aliasing, but it does not predict how important aliasing is. Our statistical approach predicts the most probable behavior of the sampled display system. This way we can easily obtain quantitative performance criteria for the overall observer-imaging system for any sampling rate.

Simple experimental observations of our environment through sampling masks indicate that with natural scenes the perceived image quality does not go through any sudden, sharp degradation as the sampling rate is changed from the "oversampled" to the "undersampled" case; as the sampling rate is continuously reduced, the perceived image quality is continuously degraded. The calculations presented in this report allow one to make quantitative estimates of this degradation as a function of sampling rate.

The key elements of our statistical approach are (a) the description of the information content of natural scenes in terms of the ensemble average power spectral density of brightness fluctuations as a function of spatial frequency; (b) the separation of a sampled image into an information containing signal component and an information degrading noise component; and (c) the calculation of the perceived image as a noisy random signal filtered by the sine wave response function of the human visual system. We found that natural scenes are well described by a power spectrum that rolls off as the inverse square of the spatial frequency. We defined three basic quality descriptors: the visual capacity that is a measure of the perceived sharpness, the perceived signal-to-noise ratio that is a measure of the perceivable number of grey levels, and an overall descriptor that combines sharpness and signal-to-noise ratio and is a measure of the total visual information transfer capacity of the display-observer system. As a consequence of the measured natural scene power spectrum, we proved that the visual capacity is not only a sharpness descriptor, but it is also proportional to the mean square perceived gradient content (expectation value of the square of brightness gradient); i.e., the visual capacity is also a measure of the actual visual information contained in a perceived noise-free image. We carried out a number of detailed numerical calculations concerning the perceived signal-to-noise ratio as a function of sampling parameters and viewing distance. Whenever applicable, preliminary simulation experiments carried out with block quantized images confirmed our theoretical predictions. Also, we showed how our formalism can be employed to optimize sampled display parameters subject to various constraints.

In all of our calculations we use a sine wave response curve of the human visual system that was obtained from one-dimensional measurements. In order

to verify that such one-dimensional response curves can be employed to describe the general, two-dimensional visual perception, we carried out measurements on a set of general, arbitrarily oriented sinusoidal stimuli that included both arbitrarily oriented one-dimensional sinusoidal gratings and two-dimensional sinusoidal brightness variations. We believe that the data presented in this report constitute the first quantitive results on the two-dimensional sine wave response of human perception. Our two-dimensional data is in general agreement with predictions based on previously obtained one-dimensional data and the assumption that linear superposition is applicable.

The report is organized in the following major parts. Section II describes the measurements and the results pertinent to the information content of natural scenes. Section III constitutes the bulk of the report and contains the mathematical description of the display-observer system applicable to sampled imaging. A second mathematical section, Section IV, treats the general two-dimensional analog display problem. Here we calculate the perceived image sharpness for displays that have anisotropic point-spread functions. This is followed by a short section describing a particularly simple method of sampled imager simulation using a real-time block quantizer. Section VI contains a detailed description of the sine wave response measurements, including both one-dimensional and two-dimensional results.

SECTION II

STATISTICAL PROPERTIES OF NATURAL SCENES

In the mathematical analysis of communication systems, it is customary to assume that the signal to be transmitted (and the noise that interferes with the faithful reception of the signal) can be treated as stationary ergodic processes in time [1]. Briefly, stationary in this sense implies that if we measure the signal (or noise) voltage, v, at times t_1 and t_2, then the statistical expectation value of the product of these two voltage readings depends only on the time difference $\tau = t_1 - t_2$. The statistical expectation value can be obtained experimentally by performing actual measurements on a large number of identical systems; i.e., on an ensemble of systems. Ergodic in this sense implies that the time and the ensemble averages are equal; the average voltage value obtained by time-averaging the signal (or noise) in a given system is equal to that obtained by averaging the simultaneous voltage readings from a large number of identical systems. For signals that describe stationary ergodic processes, the Wiener-Khinchine theorem [1] shows that the power spectral density is the Fourier transform of the auto-correlation function of the signal. One can show that the power spectral density thus defined is equal to that obtained by suitably averaging the Fourier coefficients which can be calculated through direct Fourier analysis of the signal waveform [1]. The power spectral density is a most convenient and extremely useful quantity when analyzing the performance of a given communication channel in terms of random signals.

We assume that the brightness variations of a natural scene*, as measured by any scanning aperture [2], can be considered as stationary ergodic processes.

*By "natural scenes" we mean scenes that customarily constitute the normal human environment. Our entire investigation is primarily concerned with the most probable performance of imaging systems that present a distant natural scene to a human viewer. The statistical behavior of specialized imaging systems, for example, of a microscope system used to search for flaws in periodic arrays of microcircuits will clearly be different. Most of the general mathematical formulation presented in the following sections is also applicable to the treatment of specialized imaging systems, provided that the appropriate ensemble average power spectral density is used.

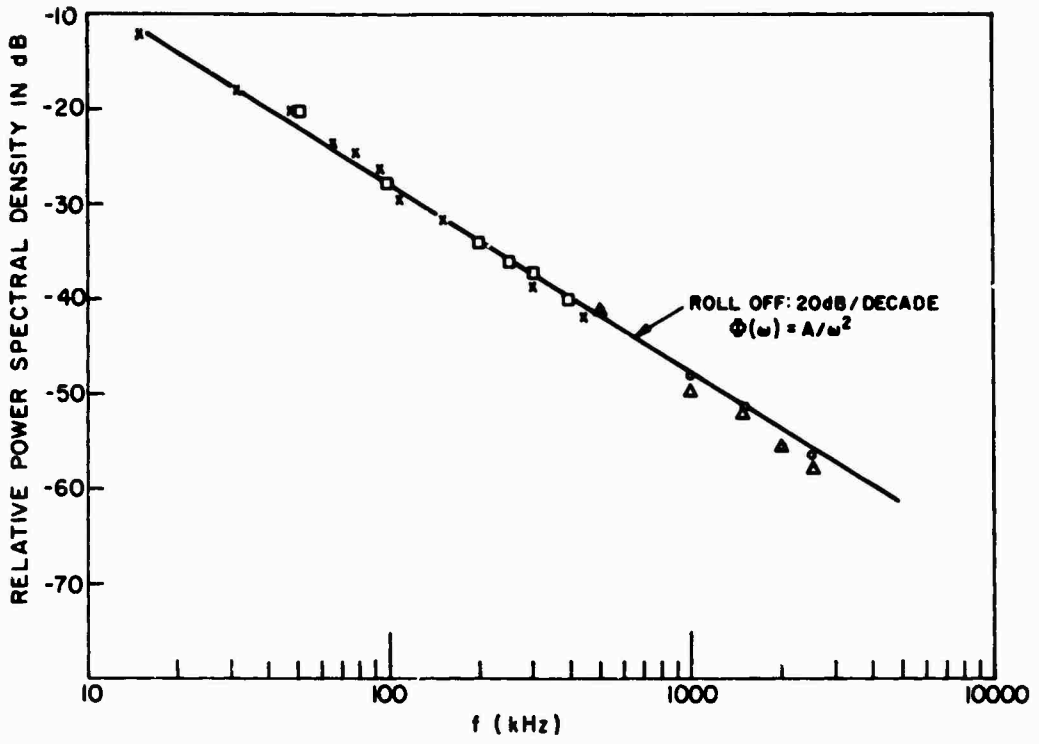
We cannot offer an explicit proof for the validity of this assumption, but we can offer two justifications for it: (a) it has been useful in connection with virtually all signals treated in communication theory, and (b) whenever possible, we compare the conclusions obtained by using this assumption with experimental observations, and unless there is a conflict, we consider it correct. The mathematical analysis presented in the following sections is based on this assumption. In order to carry out these analyses we need to know the ensemble average power spectral density for the brightness variations in natural scenes.

A convenient method to measure the desired spectral density is to analyze experimentally the spectrum of electronically transmitted brightness signals. The brightness component of commercial, off-the-air television broadcasts provides a suitable signal source for this analysis. We assume that this brightness component provides a faithful representation of the brightness variations in the original scene; i.e., we neglect the effect of system nonlinearities on the spectral power density. Even though some data are available in the literature on the spectrum of television signals [3], to our knowledge, no statistical data on a wide variety of scenes have been obtained previously.

Our experimental procedure was the following. Off-the-air television signals broadcast on VHF channels were analyzed with a commercially available spectrum analyzer. The television receiver rf, i.f., and video circuits were of sufficiently high quality that we assumed that they did not significantly alter the video spectrum. The spectrum of the U. S. standard, baseband video signals was slowly scanned at 0.4 and 4 s/MHz. A number of scans were taken at various times. In all runs the input signal amplitude was normalized by maintaining a constant sync tip to peak white video signal amplitude.

The results are shown in Fig. 1. The experimental data indicate that the ensemble average brightness variations as a function of frequency roll off smoothly at the rate of 20 dB/decade; i.e., the ensemble average power spectral density $\Phi(\omega)$ is well described by $\Phi(\omega) = B/\omega^2$, where ω is the frequency variable and B is a normalization constant that is determined by the average scene brightness.

The data points shown in Fig. 1 do not extend beyond 2.5 MHz (about 130 cycles/picture width) and below 15 kHz (about 1 cycle/picture width). On the high end, the useful range is limited by the presence of the color component in television broadcast signals. The lower limit is established by the horizontal television line scanning rate.



□ Average of 10 analyzer scans, taken on 4/11/74, 4 s/MHz scan rate.

• Average of 10 analyzer scans, taken on 3/22/74, 0.4 s/MHz scan rate.

Δ One scan, 0.4 s/MHz scan rate, taken on 3/22/74.

x Average of 3 analyzer scans, taken on 4/11/74, 4 s/MHz scan rate.

Figure 1. Statistical properties of televised scenes. The relative spectral power density as a function of frequency.

Several critical questions can be raised concerning our method of measurement. The most serious question is whether our result is simply the consequence of the television signal structure and not a measurement of scene content. The television scanning process periodically samples the scene, and the power spectrum of a periodic sequence of constant amplitude sampling gates rolls off as the inverse square of the frequency. Indeed, the spectrum of "white field" television signals (i.e., signals that are void of any scene information) do exhibit this expected power spectrum.

To check the effect of the television line structure on the spectrum of random signals with well-defined spectral properties, we performed a series of experiments in which we used electronically generated band-limited white noise with a Gaussian amplitude distribution as the "scene." Figure 2(a) shows the noise source output, and Fig. 3(a), the corresponding spectrum. In Figs. 2(a), (b), and (c), the vertical scale is signal voltage, and the horizontal scale is time with a 20-μs-per-major-division scale. In Fig. 3, the vertical scale is spectral density in dB, and the horizontal scale is a linear frequency scale, with 0 to 5 MHz full scale. The signal to be analyzed was formed by sampling the noise source output with composite video blanking and adding composite sync and a suitable white level (average brightness). In Fig. 2(b) we show the sampled noise signal and in Fig. 3(b), the corresponding spectrum; in Figs. 2(c) and 3(c) we show the composite video noise signal and the corresponding spectrum, respectively. The white level was chosen so that, when the signal was observed on an oscilloscope, the peak-to-peak noise excursions corresponded to transitions from full black to full white. The true rms voltage value of the noise signal corresponded to approximately 10 to 15% of the sync tip to full white voltage difference of the television signal thus formed. The power spectrum $\Phi(\omega)$ of this noise signal is plotted together with that of off-the-air television signals in Fig. 4. From Fig. 4 we concluded that the television sampling process does not significantly influence the power spectrum over most of the video frequency range.

A large dc component and a small random component provide another scene structure that, when formed into a television signal, could lead to an inverse square power spectrum that does not truly represent the scene content. The most obvious characteristic of such a scene is low contrast. Qualitatively, one can easily establish that natural scenes are not low contrast scenes; one need only look at one's surroundings. Quantitative results can be obtained by directly measuring local scene brightness with a spot brightness meter or by examining typical monochrome television signals in the time domain. Such measurements quickly reveal that, within the field-of-view of a typical camera setup, brightness variations well in excess of 100:1 occur; and, a typical television signal is characterized by sudden large amplitude variations rather then by small, noise-like variations superimposed on a slowly varying constant amplitude level lasting a full television line time.

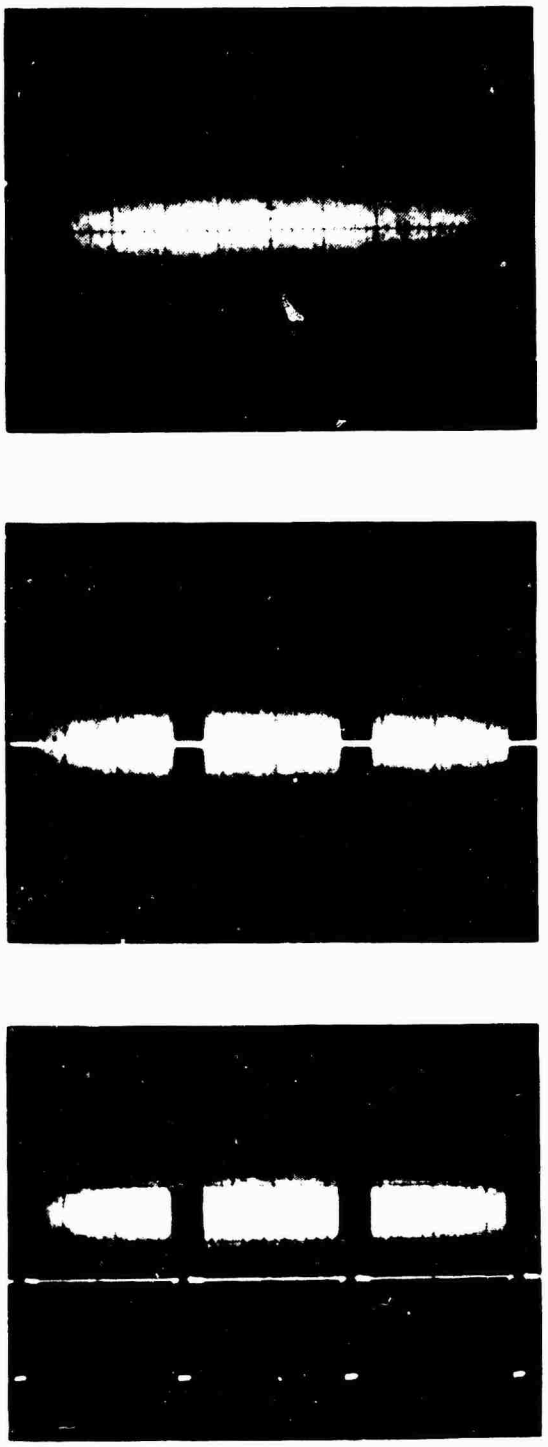


Figure 2. Band-limited Gaussian white noise signal voltage as a function of time. Vertical scale: arbitrary, consistent throughout the series; horizontal scale: 20 μs/major division. (a) Noise source output; (b) signal in (a) sampled with composite blanking signal; (c) signal in (a) formed into composite video signal.

(a)

(b)

(c)

Figure 3. Power spectral density of the signals shown in Fig. 2 as a function of frequency. Vertical scale: 10 dB per major division; horizontal scale: 0 to 5 MHz full scale. (a), (b), and (c) are the spectra corresponding to the signals of Figs. 2(a), 2(b), and 2(c), respectively.

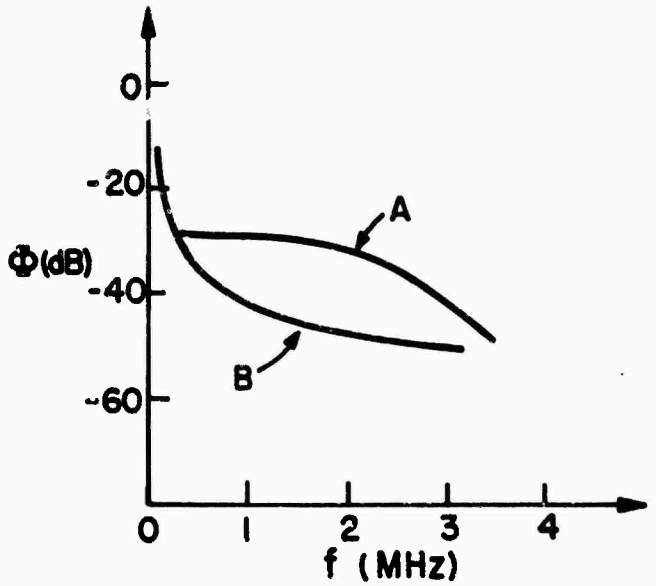


Figure 4. Power spectral density of signals as a function of frequency. Curve A: composite video signal formed from 20 Hz to 2.5 MHz band-limited Gaussian white noise. Curve B: average spectrum of off-the-air video signals, obtained from approximately 3 x 10^6 TV scan lines, black to white signal amplitude set to be approximately the same as for Curve A.

A number of additional control experiments were performed including the spectral analysis of off-the-air scenes to which test signals with known spectral properties were added, CCD camera output signals, and white fields with small amounts of additive white noise. These control experiments confirmed the validity of our main result. We conclude that our measurements using off-the-air television broadcast signals for the representation of natural scenes provide a reasonable estimate for the spatial frequency spectral content of such scenes.

In connection with the measured $1/\omega^2$ spectral behavior, one can also make the following general observations: First, note that the power spectrum of random edges (randomly spaced with random amplitudes and step heights) also rolls off as $1/\omega^2$. This is probably more than mere coincidence; it suggests very strongly that natural scenes are characterized by edge transitions. This suggestion is well supported by television practice, where good edge response is known to be a much more important quality criterion then good high-frequency response. Second, note that all cameras have a finite field-of-view and,

therefore, the integrated brightness and the integrated square of the brightness are finite. Then, it follows that the integral of the power spectral density over all frequencies must be finite, no matter how small a scanning aperture is used in measuring the brightness fluctuations. The simplest continuous high-frequency functional behavior that is integrable in the range $\omega_o \leq \omega \leq \omega_m$, with $\omega_m \to \infty$, is $1/\omega^2$. Third, that natural scenes cannot have a flat power spectrum over *any* arbitrarily low spatial frequency regime ($0 \leq \omega \leq \omega_1$) can be shown through the following simple experimental observation.* Prepare a high-resolution, large photograph of a natural scene and place a sampling mask (opaque mask with a regular array of very small openings) over it so that the photograph is viewed through the mask. When this mask-scene combination is viewed from various viewing distances, we find that, no matter how "undersampled" the image is, from sufficiently far away the only effect of the mask is a reduction of scene brightness. If there were a low-frequency regime where the scene spectrum was flat, one could always find a sufficiently low sampling rate that would result in sufficient low-frequency beats to render the sampled scene unintelligible at all viewing distances. A large number of image sampling "experiments" have been performed; for example, the images reproduced in newspapers, by CCD cameras, by digital image processing systems, and by block quantizers [4] (see Section V) are all sampled. No low-frequency sampling rate limit associated with ensembles of natural scenes has ever been found. A more concise statement of the above suggested experimental observations is that the sampling noise in sampled natural scenes can be made arbitrarily small by sufficiently increasing the viewing distance. (See the following sections for the mathematical formulation of the above statement.) As a matter of fact, it was precisely this experimental observation that motivated us to carry out the ensemble average power spectral density measurements. In our initial attempts to develop an analytical formulation of the random scene sampling problem we

*Rigorously, ω_1 cannot be allowed to approach zero, because in that case the spatial frequency bandwidth would go to zero, and, therefore, the scene would contain no information. Our experimental evidence suggests that the arguments presented here hold for $\omega_1 \gtrsim 10\ \omega_L$, where ω_L is approximately equal to the inverse picture width.

used a white spectrum and found that our results contridicted simple observations. With the measured $\Phi = B/\omega^2$ distribution, we found that our analytical results are in good agreement with experimental observations.

In the foregoing discussion we stressed the general requirement that the ensemble average power spectral density must be integrable. A function of the form B/ω^2 is not integrable at $\omega = 0$. Therefore, associated with the measured ensemble average power spectral density of brightness fluctuations there exists an effective lower scene-content cutoff frequency, ω_L. Our experimental procedure of using off-the-air television signals does not allow us to determine the precise value of ω_L experimentally; the low frequency video spectrum is dominated by the scan line structure of the video signal and not by the scene content. Nevertheless, our data indicate that the low-frequency cutoff is the order of the inverse scene width. Unless it is otherwise specified, in all of our calculations throughout this report we will use a lower cutoff angular spatial frequency $\omega_L = 2\pi/\text{image-width}$. For example, if the image extends over 10° of viewing angle at the observer, we will use a lower cutoff frequency of 0.1 cycle/degree-of-vision.

SECTION III

THEORY OF ONE-DIMENSIONAL SAMPLED DISPLAYS

A. ASSUMPTIONS AND LIMITATIONS

The approach presented here has been made possible by detailed experimental studies [5] of the one-dimensional spatial frequency response of the human visual system. An important implication of the existence of a modulation transfer function (MTF) of the visual system is that, given the perceived response to one-dimensional sine waves of known luminance modulation depth, one can, in principle, compute the perceptual response to a one-dimensional scene which consists of arbitrary luminance variations. This follows directly from Fourier's theorem, applied to one spatial dimension. Thus, it becomes possible, under certain circumstances, to treat the picture-producing device, which we call a display, and the human observer as components of an overall linear system which can be described by a set of processing parameters, transfer functions, and noise sources. In that case, one can bring to bear the mathematical apparatus of statistical communication theory to describe quantitatively the performance capabilities of the display-observer system. This is, in essence, the spirit of our approach.

The major assumptions we make in deriving the results presented here are listed below.

(1) *Linearity.* This means that both the display and observer can be treated by linear response theory, implicitly confining us to a small-signal analysis. For a linear display, we require only that there be sufficient background brightness surrounding the display to allow the perception of visual information to be treated as a perturbation. The magnitude of the adaptive luminance determines the exact form of the MTF of the human visual system. Whereas, the detailed calculations presented here were carried out using an MTF measured at very high brightness [6], our mathematical results can be applied at any level of adaptive luminance provided the proper MTF for that level is employed. On the other hand, if the display is nonlinear, we also require that the magnitude of the brightness variations on the display screen be much less than the average brightness.

(2) *Quasi-static displays.* Our analysis neglects temporal effects on the MTF of the human visual system. Therefore, strictly speaking, the results of this report apply only to scenes whose brightness variations do not change appreciably over the time required for the visual system to assume static behavior. In practice this time is on the order of several seconds [7,8].

(3) *Monochrome displays.* The presence of color, hue, and saturation variations would bring about the necessity of considering a multi-dimensional MTF of the human visual system with three frequency variables (for spatial variations of luminance, hue, and saturation) required to describe the variation of the perceived response along a single direction. Although some threshold response functions have been measured [9] for spatial modulation of color, such a multi-dimensional MTF has not been determined. Therefore, we confine ourselves here to luminance modulation of a monochrome display.

(4) *One-dimensional displays.* Although we shall be dealing here primarily with one-dimensional displays, it is possible to generalize the results of two-dimensional systems. Our first results along this line are presented in Section IV. The question of the extent of the difference between the capabilities of one- and two-dimensional descriptions is an important matter for future research.

In the sections to follow, we develop the general formulation of the one-dimensional sampled display system. We then demonstrate how the perceived response to such a display may be separated into signal and noise components, the signal containing the visual information and the noise representing a random fluctuation about the signal. Next we derive the expressions for the two important descriptors of the sampled display system - the visual capacity and the perceived signal-to-noise ratio. The former quantity is the perceptual analog of the bandwidth of communication theory and is a measure of perceived sharpness. The latter quantity is a measure of the extent to which the display noise due to sampling or other sources interferes with the perception of the visual information. Examples are given of the application of these descriptors to real and hypothetical displays. In the last section we combine the visual capacity and the perceived signal-to-noise ratio in a unified descriptor which we call the total information capacity. We demonstrate the use of this descriptor in optimizing display performance.

B. GENERAL FORMULATION OF THE ONE-DIMENSIONAL SAMPLING SCENE PROBLEM

In this section, we derive an expression for the perceived intensity pattern resulting from a particular input scene. This expression will explicitly contain the viewing distance as well as the various display parameters. It will form the basis of our subsequent quantitative description of the one-dimensional sampled display.

To begin, we assume that a particular one-dimensional signal, the input scene, is applied to the display terminals. We denote this signal by $I_o(x)$, where x is the coordinate on the display screen.* We represent the subsequent processing of the signal by the following steps:

(1) The input signal $I_o(x)$ is first passed through a filter whose (complex) transfer function is $R(\omega)$, where ω is the angular spatial frequency as measured on the display screen. The function $R(\omega)$ can have any form, subject to the normalization condition $R(0) = 1$, the condition for unity amplification of a dc signal. The filtered signal $I_F(x)$ is then given in terms of the Fourier integral

$$I_F(x) = \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi} \; R(\omega) \; \hat{I}_o(\omega) \; \exp\,(i\omega x) \tag{1}$$

where $\hat{I}_o(\omega)$ is the Fourier spectrum of the input signal $I_o(x)$.

(2) The filtered signal $I_F(x)$ is then electronically sampled within N_s equally spaced sampling gates across the display. For a display width w, the sampling frequency ω_s is defined as

$$\omega_s = 2\pi N_s / w \tag{2}$$

or, alternatively, in terms of the width $x_o = w/N_s$ of each sampling gate,

$$\omega_s = 2\pi / x_o \tag{3}$$

*For the sake of simplicity, we assume unity coordinate magnification. The extension to other magnifications is trivial.

The sampling process consists of taking the average of $I_F(x)$ within a fraction s of each sampling gate. That is, the sampled signal I_n produced by the n'th sampling gate, whose center is located at $x = nx_o$, is given by

$$I_n = (sx_o)^{-1} \int_{(n - s/2)x_o}^{(n + s/2)x_o} dx\, I_F(x) \tag{4}$$

We shall refer to the parameter s as the "sampling width." In principle, s can take on any positive value. The value s = 0 is called "delta-function sampling"; only the value of $I_F(x)$ at the center of the sampling gate contributes to I_n. On the other hand, values of s greater than unity imply that, in some regions, $I_F(x)$ contributes to more than one of the I_n. More elaborate sampling schemes can be devised and included in the formalism, but the simple averaging process, Eq. (4), is considered to be representative.

(3) Next the sampled signal is displayed on the screen by multiplying each I_n by a "printing function" $P(x - nx_o)$ to give the intensity pattern on the screen:

$$I(x) = \sum_{n=1}^{N_s} I_n\, P(x - nx_o) \tag{5}$$

The function P(x) is a property of the display. As we shall see, its particular form will be of great importance in determining display performance.

(4) The displayed pattern I(x) is perceived by an observer located a distance r from the screen. The observer is described by an MTF whose relevant frequency coordinate ν is determined at the retina and is thus measured either in cycles/radian-of-vision ($\nu = \omega r/2\pi$) or cycles/degree-of-vision ($\nu = \omega r/360$). The particular MTF $O(\nu)$ used in our calculations is shown in Fig. 5. It combines the low-frequency results of Davidson with the high-frequency measurements of Campbell and Green [6]. It is valid for high brightness levels (in excess of 100 ft-L). The MTF $O(\nu)$ is approximately

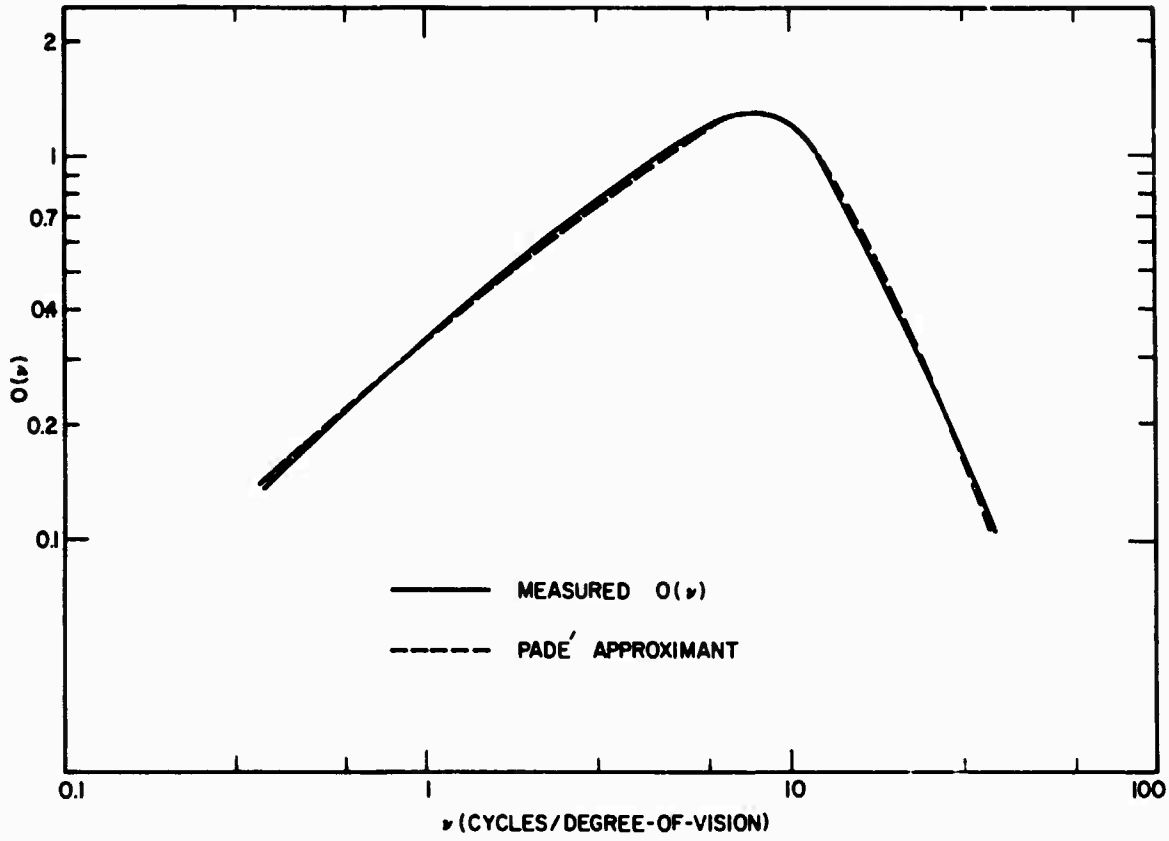


Figure 5. Modulation transfer function $O(\nu)$ of the human visual system as a function of spatial frequency ν in cycles/degree-of-vision. The solid line represents the experimental results of Davidson [6]. The dashed curve represents the approximation [11], $O(\nu) = (3.637\ \nu/\nu_o)\ [1 + 3.436\ (\nu/\nu_o) - 4.123\ (\nu/\nu_o)^2 + 2.562\ (\nu/\nu_o)^3]^{-1}$, with $\nu_o = 8$ cycles/degree-of-vision.

linear in frequency for values of ν well below the peak at 8 cycles/degree-of-vision; at high frequencies, $O(\nu)$ varies approximately as $1/\nu^2$.

We wish to calculate the intensity pattern on the screen $I(x)$ in terms of the Fourier spectrum of the input signal. First, we combine Eqs. (1) and (4) to obtain the following experssion for I_n

$$I_n = \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi}\ \text{sinc}\ (s\omega/\omega_s)\ R(\omega)\ \hat{I}_o(\omega)\ \exp\ (2\pi i n\omega/\omega_s) \qquad (6)$$

where sinc (y) is the usual function

$$\text{sinc}\ (y) = \sin(\pi y)/\pi y \qquad (7)$$

Next, we insert Eq. (6) into Eq. (5) and express $P(x - nx_o)$ as a Fourier integral to obtain

$$I(x) = \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi} \int_{-\infty}^{+\infty} \frac{d\omega'}{2\pi} \hat{P}(\omega)\ \mathrm{sinc}\ (s\omega'/\omega_s)\ R(\omega')\ \hat{I}_o(\omega')$$

$$\times \exp\ (i\omega x) \sum_{n=1}^{N_s} \exp\ [2\pi i n(\omega'-\omega)/\omega_s] \qquad (8)$$

where $\hat{P}(\omega)$ is the Fourier transform of the printing function $P(x)$. If the number of samples N_s is very large, we can evaluate the summation over n in Eq. (8) by making use of the mathematical identity

$$\sum_{n=-\infty}^{+\infty} \exp\ (iny) = 2\pi\delta(y + 2\pi m);\ m = 0, \pm 1, \pm 2, \ldots \qquad (9)$$

where $\delta(y)$ is the Dirac delta-function. Employing Eq. (9), we perform the integration over ω' in Eq. (8) to obtain

$$I(x) = \sum_{m=-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi}\ x_o^{-1}\ \hat{P}(\omega)\mathrm{sinc}[s(\omega-m\omega_s)/\omega_s]\ R(\omega-m\omega_s)$$

$$\times \hat{I}_o(\omega-m\omega_s)\ \exp\ (i\omega x) \qquad (10)$$

At this point, we introduce a normalization constraint on $\hat{P}(\omega)$. Since we will want to compare displays with equal total light output, we insist that, for a unit dc input, the integrated intensity $\int_{-\frac{1}{2}w}^{+\frac{1}{2}w} dx\ I(x) = w$ regardless of the form of $P(x)$. Employing Eq. (10), with $\hat{I}_o(\omega) = 2\pi\delta(\omega)$, we readily obtain the condition $\hat{P}(0) = x_o$, again valid for $N_s >> 1$. Therefore, the factor $x_o^{-1}\hat{P}(\omega)$ in Eq. (10) may be replaced by $\hat{P}(\omega)/\hat{P}(0)$. Instead we choose to renormalize $\hat{P}(\omega)$ such that

$$\hat{\Pi}(\omega) = x_o^{-1}\,\hat{P}(\omega) = \hat{P}(\omega)/\hat{P}(0) \tag{11}$$

According to step (4) above, the intensity pattern Eq. (10) is transferred to the perceptual level by means of the MTF of the human visual system $O(\nu)$. Since Eq. (10) is in the form of a Fourier integral, the perceived pattern $E(x)$ is obtained from this equation by simply multiplying* the integrand of Eq. (10) by the function $O(\omega r/2\pi)$. Thus, using Eq. (11), we have the result for $E(x)$

$$E(x) = \sum_{m=-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi}\, O(\omega r/2\pi)\, \hat{\Pi}(\omega)\,\mathrm{sinc}[s(\omega - m\omega_s)/\omega_s]$$

$$\times\, R(\omega - m\omega_s)\, \hat{I}_o(\omega - m\omega_s)\, \exp(i\omega x) \tag{12}$$

Eq. (12) is the major result of this section. It expresses the perceived pattern in terms of the viewing distance, the various filter functions and parameters of the display system, and the spectrum of the input scene.

Equation (12) shows that the perceived intensity can be regarded as the sum of various contributions corresponding to a displacement of the spectral content of the input scene $\hat{I}_o(\omega)$ by multiples of the sampling frequency. The situation is shown schematically in Fig. 6, where the contributions of specific m-values to the total spectral content are plotted as a function of frequency. Two distinct cases are shown, depending on the relative values of the sampling frequency ω_s and the maximum frequency ω_M for which the filter function $R(\omega)$ is non-zero.

As shown in Fig. 6(a), if $\omega_s > 2\omega_M$, the contributions to the integrand of Eq. (12) from the various m-values do not overlap in frequency. Such a display obeys the so-called Nyquist criterion [10], and we call displays of this type "oversampled". Suppose now that the MTF of the human visual system could be replaced by unity for all frequencies whose magnitude is less than a value ω_c such that $\omega_M \leq \omega_c \leq (\omega_s - \omega_M)$. Then, using Eq. (12), it is easy to show

*We neglect any angular factors that arise from such sources as oblique angle viewing and Lambertian light emission from the display screen. To the extent that these effects only alter the total brightness of the display, they can be included in the function $\hat{I}_o(\omega)$.

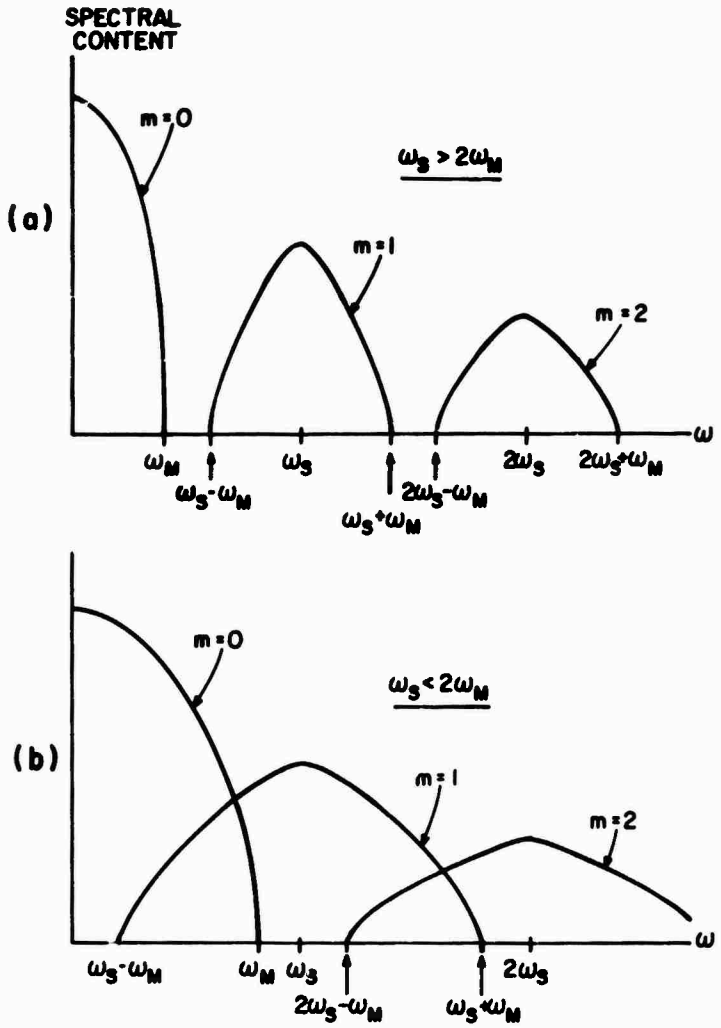


Figure 6. Schematic representation of the spectral content of a sampled display as a function of spatial frequency: (a) a display that obeys the Nyquist criterion $\omega_S > 2\omega_M$; (b) a display for which $\omega_M < \omega_S < 2\omega_M$.

that, if we take s = 0 (delta-function sampling), $\hat{\Pi}(\omega) = 1$ (delta-function printing function), and $R(\omega) = 0$ for $|\omega| > \omega_M$, the function $E(x)$ is identical to one that is produced by merely sending the input scene through the filter function $R(\omega)$ without undergoing the subsequent sampling process. This result is equivalent to the well-known Shannon theorem [10] for communication systems that obey the Nyquist criterion. Of course, the MTF of the human system cannot be represented by such a simple function. Thus, although the display obeys the Nyquist criterion, it is not possible to transfer to the perceptual level an exact replica of the original band-limited signal. However, the Nyquist

criterion, per se, is irrelevant for the human observer viewing a sampled display with $\omega_s > \omega_M$. If we accept the concept of an effective cutoff frequency ν_c of the human visual system, we can increase the viewing distance until $\nu_c < (\omega_s - \omega_M)r/2\pi$. At such viewing distances, the viewer would be unable to perceive the difference between the sampled display and an analog display with an overall band-limited MTF $R(\omega)$. We are thus led to a theorem appropriate for sampled displays: *For display systems with* $\omega_s > \omega_M$, *there exists a range of viewing distances* $r > 2\pi\nu_c/(\omega_s - \omega_m)$ *such that, with delta-function sampling and printing, the display is indistinguishable from an analog display with the same band-limiting characteristics, viewed at the same distance.* (In other words, we require only one sample per wavelength instead of the customary two samples per wavelength, as set by the general sampling theorem [10].) As an example, we estimate the required viewing distance for a hypothetical display sampled at the Nyquist rate $\omega_s = 2\omega_M$. Then, taking* $\nu_c \tilde{=} 60$ cycles/degree-of-vision = 3440 cycles/radian-of-vision, we have $r/w \gtrsim 6900/N_s$. For 700 samples, the theorem says that such a display would be indistinguishable from the equivalent analog display at viewing distances greater than about 10 picture widths.

In Fig. 6(b), we indicate the situation for displays that have $\omega_s < 2\omega_M$ and thus do not obey the Nyquist criterion. We call displays of this type "undersampled". As can be seen from the figure, the contributions to the integrand of Eq. (12) from the various m-values overlap in frequency. Therefore, according to Shannon [10], it is impossible to reconstruct the original band-limited signal $I_F(x)$ no matter what form of sampling or printing function is chosen and no matter what form we take for the MTF of the human visual system. However, according to the theorem stated above, as long as $\omega_s > \omega_M$, we can devise a sampling and printing technique such that an observer can position himself sufficiently far away for him to be unaware of the difference between the sampled display and an analog display with an overall MTF $R(\omega)$. On the other hand, if $\omega_s < \omega_M$, it is always possible to construct a scene such that terms in Eq. (12) with $m \neq 0$ can be perceived no matter what the viewing distance. The condition $\omega_s < \omega_M$ is simply the condition that allows Moiré beats to be generated at dc.

*This value of ν_c corresponds to an acuity of approximately 1 min of angle [13].

C. SEPARATION OF THE PERCEIVED INTENSITY INTO SIGNAL AND NOISE COMPONENTS

In this section, we employ general arguments regarding the invariance and statistical properties of E(x) in order to justify the separation of E(x) into signal and noise components. This separation is essential to the derivation of the descriptors given later in this report.

As can be seen from Eq. (12), the total perceived response can be written as the sum of contributions from the various m-values:

$$E(x) = \sum_{m=-\infty}^{+\infty} E_m(x) \tag{13}$$

where

$$E_m(x) = \int_{m\omega_s-\omega_M}^{m\omega_s+\omega_M} \frac{d\omega}{2\pi}\, O(\omega r/2\pi)\, \hat{\Pi}(\omega) \operatorname{sinc}[s(\omega-m\omega_s)/\omega_s] \times R(\omega-m\omega_s)\, \hat{I}_o(\omega-m\omega_s) \exp(i\omega x) \tag{14}$$

In Eq. (14) we have explicitly included the integration limits imposed by the finite passband associated with $R(\omega)$. We now separate the contributions to E(x) into a signal part $E_S(x)$ and a noise part $E_N(x)$ according to the following prescription:

$$E(x) = E_S(x) + E_N(x) \tag{15}$$

where

$$E_S(x) = E_o(x) = \int_{-\omega_M}^{+\omega_M} \frac{d\omega}{2\pi}\, O(\omega r/2\pi)\, \hat{\Pi}(\omega) \operatorname{sinc}(s\omega/\omega_s) \times R(\omega)\, \hat{I}_o(\omega) \exp(i\omega x) \tag{16}$$

$$E_N(x) = \sum_{m\neq 0}^{+\infty} E_m(x) = \sum_{m\neq 0}^{+\infty} \int_{m\omega_s-\omega_M}^{m\omega_s+\omega_M} \frac{d\omega}{2\pi}\, O(\omega r/2\pi)\, \hat{\Pi}(\omega)$$

$$\times \operatorname{sinc}[s(\omega-m\omega_s)/\omega_s]\, R(\omega-m\omega_s)\, \hat{I}_o(\omega-m\omega_s)\, \exp(i\omega x) \qquad (17)$$

Now, the question arises as to what extent the quantities $E_S(x)$ and $E_N(x)$ truly represent signal and noise contributions, respectively, to the total perceived pattern. At the outset, one may legitimately argue that true noise is unpredictable, since it arises from random processes, so that the quantity defined in Eq. (17), being completely determined by the input spectrum and the various operations performed by the display, does not represent true noise. In fact, the quantity E_N represents a kind of distortion in which any input frequency ω produces output amplitudes at $(\omega + m\omega_s)$ for all integer values of m. However, in what follows, we show that, in a statistical sense, the noise part $E_N(x)$ of the total perceived intensity E(x) is uncorrelated with the input scene $I_o(x)$, and, thus, $E_N(x)$ does possess a fundamental property of noise.

THEOREM: If the ensemble of input scenes is translationally invariant (i.e., stationary in a statistical sense) $E_N(x)$ *and* $I_o(x)$ *are uncorrelated.*

To prove the theorem, we will calculate the quantity $\langle\overline{E(x)I_o(x)}\rangle$, where the bar denotes an average over x, and the bracket symbol stands for an average over an ensemble of many scenes. From Eq. (12) for E(x), we have

$$\langle\overline{E(x)I_o(x)}\rangle = (1/w) \int_{-\frac{1}{2}w}^{+\frac{1}{2}w} dx \sum_{m=-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi} \int_{-\infty}^{+\infty} \frac{d\omega'}{2\pi}\, O(\omega r/2\pi)$$

$$\times \hat{\Pi}(\omega) \operatorname{sinc}[s(\omega-m\omega_s)/\omega_s]\, R(\omega-m\omega_s)$$

$$\times \langle \hat{I}_o(\omega-m\omega_s)\, \hat{I}_o(\omega')\rangle \exp[i(\omega+\omega')x] \qquad (18)$$

In order to obtain the desired result, we must prove the following lemma.

LEMMA: If the ensemble of scenes is translationally invariant, $\langle \hat{I}_o(\omega)\hat{I}_o^*(\omega')\rangle$ *is of the form* $2\pi\Phi(\omega)\delta(\omega-\omega')$.

PROOF: We write the autocorrelation function for the ensemble of input scenes in terms of the spectra of these scenes:

$$\langle I_o(x)\ I_o(x+L)\rangle = \int_{-\infty}^{+\infty}\frac{d\omega}{2\pi}\int_{-\infty}^{+\infty}\frac{d\omega'}{2\pi}\langle \hat{I}_o(\omega)\ \hat{I}_o(\omega')\rangle$$

$$\times \exp[i(\omega+\omega')x + i\omega' L]$$

$$= \int_{-\infty}^{+\infty}\frac{d\omega}{2\pi}\int_{-\infty}^{+\infty}\frac{d\omega'}{2\pi}\langle \hat{I}_o(\omega)\ \hat{I}_o^{*}(\omega')\rangle$$

$$\times \exp[i(\omega-\omega')x - i\omega' L] \qquad (19)$$

where we have used the relation $\hat{I}_o(\omega) = I_o(-\omega)$, a consequence of the reality of $I_o(x)$. From Eq. (19), it is seen that, for an arbitrary $\hat{I}_o(\omega)$, the only way the autocorrelation function can be independent of x is for

$$\langle \hat{I}_o(\omega)\ \hat{I}_o^{*}(\omega')\rangle = 2\pi\ \Phi(\omega)\delta(\omega'-\omega) \qquad (20)$$

where $\Phi(\omega)$ is a real function, the power spectrum of the ensemble of scenes. Thus, the lemma is proved, and Eq. (19) reduces to the Wiener-Khinchin theorem [1]; i.e., the autocorrelation function is the Fourier transform of the power spectrum.

Returning now to Eq. (18), use of the lemma gives rise to

$$\overline{\langle E(x)I_o(x)\rangle} = (1/w)\int_{-\frac{1}{2}w}^{+\frac{1}{2}w} dx \sum_{m=-\infty}^{+\infty} \int_{-\infty}^{+\infty}\frac{d\omega}{2\pi}\, O(\omega r/2\pi)\ \hat{\Pi}(\omega)$$

$$\times \operatorname{sinc}[s(\omega-m\omega_s)/\omega_s]\ R(\omega-m\omega_s)$$

$$\times \Phi(\omega-m\omega_s)\ \exp(im\omega_s x) \qquad (21)$$

For a very large display, the integration over x gives a contribution only from the term m = 0, since all terms with m≠0 are pure oscillatory in x and average to zero. We thus have the desired result

$$\overline{\langle E(x)I_o(x)\rangle} = \overline{\langle E_S(x)\ I_o(x)\rangle}$$

$$\overline{\langle E_N(x)I_o(x)\rangle} = 0 \qquad (22)$$

where $E_S(x)$ and $E_N(x)$ are the signal and noise contributions, respectively, as defined in Eqs. (16) and (17).

The above theorem shows that, on the average, $E_N(x)$ does not contribute to the correlation function $\overline{\langle E(x)I_o(x)\rangle}$. This means that we may regard E(x) as consisting of a correlated part $E_S(x)$ plus a fluctuating part $E_N(x)$, where the average value of the fluctuation about $E_S(x)$ vanishes. Thus, even though $E_N(x)$ is not the product of the kind of random process we normally associate with noise, it does display the fundamental property of noise of being uncorrelated with the input. Simply put, this means that if we look at a particular place on a display screen on which an arbitrary scene has been produced, there is no reason to expect that noise will either add to or subtract from the perceived signal. The expectation value of the product of the input signal and the perceived noise for an arbitrary input scene should vanish. Indeed, Eq. (22) shows that the separation $E(x) = [E_S(x) + E_N(x)]$, where $E_S(x)$ and $E_N(x)$ are given in Eqs. (16) and (17), satisfies this criterion.

Furthermore, we can show that $E_S(x)$ and $E_N(x)$ have significantly different invariance properties under an arbitrary translation of the input image. Consider a displacement δx of the input scene. From the form of $E_S(x)$ given in Eq. (16), it is obvious that $E_S(x)$ can be regarded as being generated by an effective display MTF $R_{eff}(\omega)$ such that

$$R_{eff}(\omega) = \hat{\Pi}(\omega)\ \mathrm{sinc}(s\omega/\omega_s)\ R(\omega) \qquad (23)$$

Thus, if $I_o(x) \rightarrow I_o(x + \delta x)$, we must have $E_S(x) \rightarrow E_S(x + \delta x)$; we obtain the same signal but shifted by the same distance as was the input. However, from the form of $E_N(x)$ given in Eq. (17), it is easy to show that, for arbitrary δx and $\hat{I}_o(\omega)$, none of the $E_m(x)$ comprising $E_N(x)$ transforms to $E_m(x + \delta x)$.

Indeed, we find that, under the translation $I_o(x) \rightarrow I_o(x+\delta x)$, $E_m(x) \rightarrow E_m(x+\delta x)\exp(im\omega_s \delta x)$. Thus, $E_N(x)$ is, in general, altered in a complicated way by the translation. If $E_N(x)$ arose from random processes not associated with the scene content itself, we would not expect $E_N(x)$ to depend on the "phase" of the input signal. Nevertheless, a fundamental property of any communication system is that, for an arbitrary displacement of the input stimulus, the output message should also translate uniformly to give the same signal but displaced by the same amount as the input. No distortions of the signal should be introduced by a decision to send a message earlier or later than was originally intended. The quantity $E_S(x)$, defined in Eq. (16), does indeed meet this requirement, whereas $E_N(x)$, defined in Eq. (17), does not display this translational invariance.

The above properties of $E_S(x)$ and $E_N(x)$ show that our method of separating the perceived intensity into signal and noise components is justified. The quantity $E_S(x)$ is correlated with the input scene. It has the same translation properties as the input scene, and so it can be described by an effective MTF, given in Eq. (23). The quantity $E_N(x)$ is uncorrelated with the input. It does not respond to translations of the input scene in the same way as $E_S(x)$ but rather gives rise to a random error that depends on the input scene, its phase with respect to the sampling locations, and the parameters of the sampling process.

D. TWO DESCRIPTORS: THE VISUAL CAPACITY AND THE PERCEIVED SIGNAL-TO-NOISE RATIO

With the formalism developed in the preceding sections, we are in a position to calculate two fundamental descriptors for the one-dimensional sampled display. In the spirit of our treatment of the display-observer system as a linear communication channel, it is natural to consider quantities that are analogous to the well-known quality descriptors for such systems: bandwidth and signal-to-noise ratio. In the discussion below, we develop the expressions for the perceptual analog of these important quantities. The quantity analogous to the bandwidth is the visual capacity, a measure of edge discrimination ability in the absence of noise. The perceived signal-to-noise ratio is a determination of the relative amount of signal transmitted to the perceptual level compared with the corresponding noise.

1. The Visual Capacity

In Ref. [11], two of the authors of this report developed a descriptor that is a direct measure of the number of fully resolvable edge transitions that can be perceived across a noiseless, analog display of width w. This descriptor, called the total visual capacity and denoted by C_v^T, is computed from the perceived width x_e of an edge transition according to the relation [11]

$$C_v^T = w/x_e \tag{24}$$

The edge width x_e, in turn, is computed from the integral of the square of the magnitude of the overall MTF of the display-observer system [11]

$$1/x_e = (1/\pi) \int_0^\infty d\omega \; O^2(\omega r/2\pi) \; |R_o(\omega)|^2 \tag{25}$$

Here $R_o(\omega)$ represents the overall MTF of the display system, including band limitations of the input as well as the limitations of the display device itself. Combining Eqs. (24) and (25), we have, making the dependence of C_v^T on viewing distance explicit,

$$C_v^T(r) = (w/\pi) \int_0^\infty d\omega \; O^2(\omega r/2\pi) \; |R_o(\omega)|^2 \tag{26}$$

As discussed in Ref. [11], C_v^T can be regarded as the information capacity of a noiseless two-level communication channel. It is also a generalization to the perceptual level of the noise equivalent bandwidth that was proposed by O. H. Schade [2] as an image quality descriptor. Schade argued that the specification of the limiting resolution alone is inadequate as a descriptor. Instead, the noise equivalent bandwidth, which weights a given frequency according to the square of the MTF at that frequency, provides a better overall quantitative measure of the effective bandwidth of the system. The authors of Ref. [11] showed that, when generalized to the perceptual level, the noise equivalent bandwidth is actually a measure of perceived edge sharpness.

The usefulness of the visual capacity has been demonstrated [11] in the prediction of such quantities as the optimum viewing distance and the relative performance of various display systems at any viewing distance, all within the restriction that we confine ourselves to noiseless analog systems. In order to extend Eq. (26) to the case of sampled displays, we first recall that, according to the discussion of the previous section, the perceived response to a scene imaged on a sampled display may be regarded as the sum of the signal and noise contributions. The signal contribution is described by an effective MTF $R_{eff}(\omega)$ for sampled displays, given in Eq. (23), while the noise contribution is uncorrelated with the ensemble of input scenes. Therefore, the visual capacity for sampled displays is obtained from Eq. (26) by substituting $R_{eff}(\omega)$ for the overall MTF $R_o(\omega)$ of the analog display

$$C_v^T(r) = (w/\pi) \int_0^\infty d\omega\, O^2(\omega r/2\pi)\, |\hat{\Pi}(\omega)|^2\, \mathrm{sinc}^2(s\omega/\omega_s) |R(\omega)|^2 \qquad (27)$$

Sampled Displays

It should be emphasized that Eq. (27) for C_v^T is a measure of edge discrimination ability, ignoring the effect of the sampling noise $E_N(x)$. As such, it represents the effective bandwidth of the display-observer communication system. In our formulation, the deleterious effect of the noise is not included in this bandwidth but rather is included in the perceived signal-to-noise ratio, to be discussed later.

a. Properties of the Visual Capacity. - The general properties of C_v^T for analog display systems have been discussed in Ref. [11]. We note here that, for the case $R(\omega) = R_o(\omega)$, a comparison of Eqs. (26) and (27) for the C_v^T of analog and sampled display systems, respectively, shows that *sampling always degrades the perceived sharpness of the equivalent analog display.* This observation follows immediately from the fact that both $|\hat{\Pi}(\omega)|$ and $\mathrm{sinc}(s\omega/\omega_s)$ are always less than or equal to unity [12]. It is clear that the visual capacity of a sampled display is increased by raising the sampling frequency, or, alternatively, by employing a very small sampling width s and a very narrow printing function P(x). Indeed in the limit of delta-function sampling $s = 0$ and delta-function printing $\hat{\Pi}(\omega) = 1$, we recover the visual capacity of an equivalent analog display with a band-limited input characteristic $R(\omega)$ and a

perfect writing beam. This conclusion may appear to be somewhat surprising since it does not contain an explicit statement about the value of N_s; the only requirement is that $s = p = 0$. There is, in fact, an implicit requirement that $N_s \gg 1$, because our formalism is based on the assumption that the number of samples is sufficiently large to allow the replacement of the summation over the finite number of sampling cells by an infinite sum [see Eqs. (8) and (9)]. Furthermore, it is important to keep in mind that the sampled image can be separated into a signal and a sampling noise component only in a statistical sense, and that this separation is essential for the definition of $R_{eff}(\omega)$ and, therefore, for the calculation of C_v^T. The reader should grasp the essential difference between $R_{eff}(\omega)$ and the MTF of an ordinary analog display. In the case of an analog display, the reproduced contrast ratio for any given sinusoidal input of frequency ω is explicitly specified by the modulation transfer function $R(\omega)$; in the general sampled display case, $R_{eff}(\omega)$ by itself does not specify the reproduced contrast because the displayed signal is always accompanied by the sampling noise. However, as shown in Section III.C, $R_{eff}(\omega)$ does indeed give the average response of the display, so that Eq. (27) for C_v^T represents the number of perceivable edges after suitable averaging cancels the sampling noise. The effect of the noise will be included separately in the calculation of the perceived signal-to-noise ratio. We shall see later that, whereas decreasing the sampling width and the width of the printing function improves the visual capacity, such a practice degrades the perceived signal-to-noise ratio. The optimum strategy can only be determined by considering both visual capacity and perceived signal-to-noise ratio.

The asymptotic behavior of C_v^T at very small and very large viewing distances can be obtained by extending the technique employed in Ref. [11]. Each of the four functions appearing in the integrand of Eq. (27) have a characteristic range in ω-space. These ranges are:

Range of $O(\omega r/2\pi) \sim 2\pi\nu_o/r$
Range of $|R(\omega)| \sim \omega_M$
Range of $\text{sinc}\,(s\omega/\omega_s) \sim \omega_s/s$
Range of $|\hat{\Pi}(\omega)| \sim \omega_s/p$.

Here ν_o is the retinal frequency for which $O(\nu)$ has a maximum; from Fig. 5, it is seen that ν_o = 8 cycles/degree-of-vision = 458 cycles/radian-of-vision. The quantity p defines the effective dimensionless width of the printing function P(x); i.e., the quantity px_o is the range of P(x).

The far-field viewing distance limit of $C_v^T(r)$ is obtained for such large values of r that the range of $O(\omega r/2\pi)$ is much less than those of the other three functions appearing in the integrand of Eq. (27). In that case, $R(\omega)$, $\hat{\Pi}(\omega)$, and $\text{sinc}(s\omega/\omega_s)$ may be replaced by unity in Eq. (27) to give

$$C_v^T(r) = w/\theta_c^\infty r \tag{28}$$

where

$$1/\theta_e^\infty = 2 \int_0^\infty d\nu \, O^2(\nu) \tag{29}$$

is the perceived angular width of a single edge transition, as reproduced by a perfect display. As discussed in Ref. [11], we have adopted a scale for $O(\nu)$ that gives a maximum value of the total visual capacity equal to the number of TV lines $N_{TV} = \omega_M w/\pi$ that can be produced by an analog display with a flat response $R(\omega) = 1$ below the maximum frequency ω_M. This procedure establishes the value [11]

$$\theta_e^\infty = 1.84 \text{ min of angle} \tag{30}$$

for the perceived angular width of a perfect edge. This value lies within the range found in early measurements of visual acuity on line patterns [13]. However, the numerical value of θ_e^∞ given here has no fundamental perceptual significance; it is used only to establish a convenient absolute scale for C_v^T that allows easy appreciation of any visual capacity value in terms of a corresponding ideal TV system. Combining Eqs. (29) and (30) gives the following simple law for the visual capacity at very large viewing distances:

$$C_v^T(r) = 1870 \, w/r; \text{ for } r/2\pi\nu_o >> 1/\omega_M, \, p/\omega_s, \, s/\omega_s \tag{31}$$

In the limit of very small viewing distances such that the range of the integrand of Eq. (27) is not affected by the high-frequency rolloff of $O(\omega r/2\pi)$, the function $O(\omega r/2\pi)$ can be replaced by its low-frequency, linear asymptote [11]

$$\lim_{\nu \to 0} O(\nu) = 3.637\ \nu/\nu_o \tag{32}$$

We have then

$$C_v^T(r) = 4.21\ w(r/2\pi\nu_o)^2 \int_0^\infty d\omega\ \omega^2\ |\hat{\Pi}(\omega)|^2\ \mathrm{sinc}^2(s\omega/\omega_s)\,|R(\omega)|^2; \tag{33}$$

$$\text{for } r/2\pi\nu_o \ll \text{Largest of } 1/\omega_M,\ p/\omega_s,\ s/\omega_s$$

From Eq. (33) it is seen that, in near-field viewing, $C_v^T(r)$ increases as the square of the viewing distance, with the magnitude determined by the second moment of the effective display MTF. The occurrence of the second moment emphasizes the high-frequency response of the display system. This result can be understood by considering that, when we view an image from such small distances that all spatial frequencies produced by the display lie below the peak in the response of the human visual system, we would be quite capable of perceiving detail that is, in fact, not being produced by the display. In this case, visual capacity is increased either by increasing the high-frequency response of the display or by moving further away so as to bring the peak of the response of the human visual system within the effective passband of the display.

Since $C_v^T(r)$ rises as r^2 for small viewing distances and falls off as $1/r$ at large viewing distances, $C_v^T(r)$ must achieve a maximum value at some intermediate viewing distance. One anticipates that the maximum value of $C_v^T(r)$ is achieved at the viewing distance for which the peak of the MTF of the human visual system corresponds approximately to the rolloff frequency of the effective display MTF $R_{eff}(\omega)$. Thus, we expect that the peak value of the visual capacity should occur at a viewing distance r_p such that

$$r_p/2\pi\nu_o \sim \text{Largest of } 1/\omega_M,\ p/\omega_s,\ s/\omega_s \tag{34}$$

Eq. (34) simply says that the viewing distance for maximum $C_v^T(r)$ is determined by matching the peak of the eye's sensitivity curve to the effective bandwidth of the display.

b. Calculated Examples. - An example of the degradation of the visual capacity due to sampling is shown in Fig. 7. The curves shown in the figure were calculated from Eq. (27) for hypothetical sampled displays with the indicated values of the number of samples N_s and with a flat pre-sampling filter function $R(\omega) = 1$ up to the maximum (cutoff) frequency ω_M. The value of ω_M was chosen to give a limiting resolution of $N_{TV} = \omega_M w/\pi = 500$ lines. We have taken full-width sampling $s = 1$ and a printing function that corresponds to constant illumination of each sampling aperture. Thus, we have $P(x) = 1$ for $|x| \leq \frac{1}{2} x_o$ and $P(x) = 0$ for $|x| > \frac{1}{2} x_o$, giving $\hat{\Pi}(\omega) = \text{sinc}\,(\omega/\omega_s)$. The curve labeled "$N_s = \infty$" corresponds to the 500-line analog system, so that $C_v^T(r)$ has a maximum value of 500 transitions, as it should. As N_s is decreased, $C_v^T(r)$ decreases at all viewing distances, and r_p is shifted to higher values from its value of about 2.7 picture widths for the analog limit. The calculated drop of the visual capacity is gradual at first but becomes severe for $N_s \lesssim 1000$. We notice that, since

$$N_{TV}/N_s = 2\omega_M/\omega_s \tag{35}$$

when N_s has been reduced to 500 samples, the display is being sampled at precisely the Nyquist rate $\omega_s = 2\omega_M$. At this point, the peak visual capacity is equivalent to only a 280-line analog display, and r_p has been shifted to about 3.7 picture widths. Further reductions of N_s would cause an even greater degradation of $C_v^T(r)$.

In Fig. 8, we have used Eqs. (26) and (27) to compare $C_v^T(r)$ for the horizontal and vertical directions of a real television display device consisting of a wide-angle, small-screen color kinescope. This is of particular interest because this display is a hybrid system; the image-forming capability in the horizontal direction is that of an analog scanning display, whereas the image-forming method in the vertical direction is that of a sampled display. In the calculations, we used a measured* kinescope MTF for a single electron

*E. M. Herold, private communication.

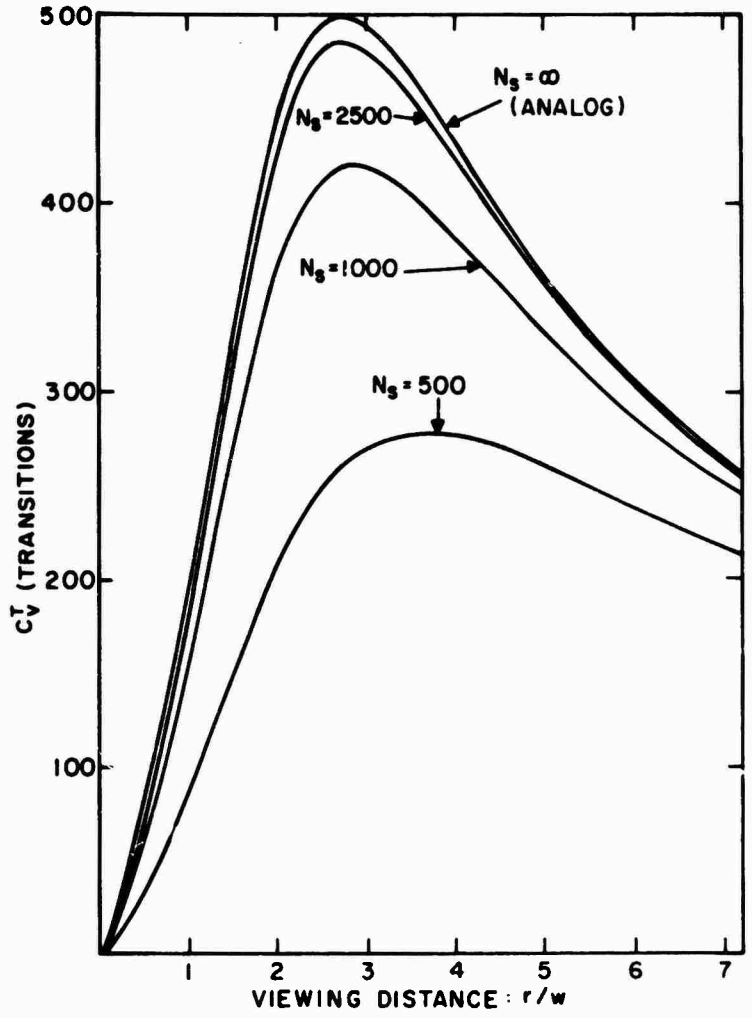


Figure 7. Visual capacity C_V^T as a function of viewing distance in picture widths for a display system with a flat passband giving a limiting resolution N_{TV} = 500 lines. The number of samples N_S for each of the curves is indicated in the figure.

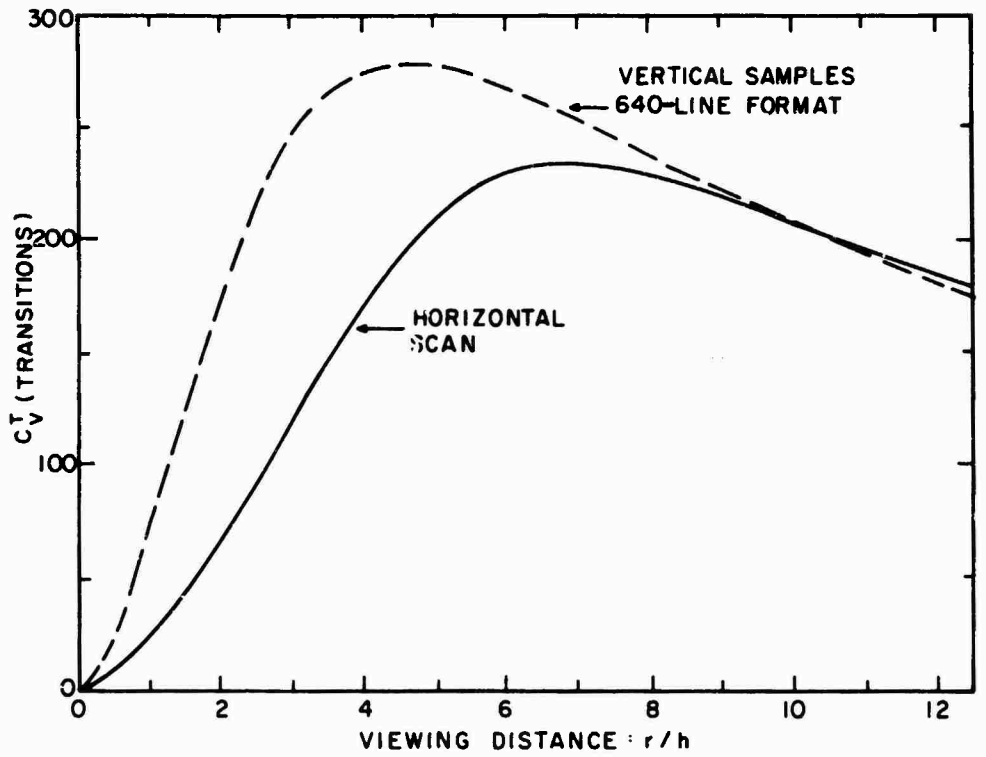


Figure 8. Visual capacity C_V^T as a function of viewing distance in picture heights for a real television display device. The C_V^T for both vertical and horizontal directions are shown. An N_S = 640-line format has been employed for the vertical direction in order to compensate for the 4:3 aspect ratio of the display screen.

gun in order to obtain the effective MTF's for the horizontal and vertical directions. The beam current level in the kinescope at which the MTF's were determined corresponds to approximately 100-ft-L brightness. The contribution to $R_o(\omega)$ of the video response function limiting the range of frequencies available for the horizontal signals was taken from typical measured* circuit characteristics. For the vertical direction, we assumed extreme undersampling and, therefore, took $\omega_M = \infty$. In addition, we used the value s = 1, corresponding to full-width sampling. In order to effect a direct comparison of the horizontal and vertical $C_V^T(r)$, we compensated for the 4:3 aspect ratio of the display screen by computing $C_V^T(r)$ for the vertical direction using a 640-sample format rather than the actual number, 480 samples, contained in one picture height. However, the viewing distance coordinate is measured in picture

*J. R. van Raalte and W. G. Gibson, private communication.

heights h, as is customary in the television industry. In performing the calculations, we assumed an isotropic $O(\nu)$, a reasonably good approximation, as shown elsewhere in this report. It is seen from Fig. 8 that the sharpness capability of the horizontal and vertical directions are only fairly well matched in this case. The peak visual capacity for the horizontal direction is about 18% less than that for the vertical direction, and this peak occurs at a considerably larger viewing distance, about 7h compared with 4.7h for the vertical direction. The major cause of the calculated anisotropy is the incomplete utilization of the bandwidth available for the transmission of luminance signals for horizontal display.

In Fig. 9, we illustrate the effect of varying the width of the printing function on the $C_v^T(r)$ for the vertical (sampled) direction of the television display described above. Here we use a 480-sample format to compute the actual visual capacity for the display of height h. The parameter p represents the effective width of the printing function. In this case, P(x) was approximately Gaussian, and we have defined p as the distance, in units of x_o, at which the printing function falls to 1/e of its maximum value. The particular value p = 0.642 was the value observed experimentally, and the other values of p for which curves are shown represent successive incremental changes of about 20% from the experimental value. It is seen that, by increasing p, we degrade the visual capacity and increase r_p, whereas the visual capacity is enhanced and r_p moves to smaller values as p is decreased from the experimental value. This behavior is an example of the general observation that narrow width printing and sampling functions always enhance $C_v^T(r)$. In the limit $p \to 0$, we calculate that the maximum value of $C_v^T(r) = 375$ transitions at a viewing distance $r_p/h = 2.7$. This result shows that an s = 1 sampled display system with N_s samples operating in the extreme undersampled limit has a maximum edge discrimination ability equivalent to an analog display with a flat passband and limiting resolution given by $N_{TV} = (375/480)\ N_s = 0.78N_s$.

c. The Perceived Gradient Content and its Relation to the Visual Capacity. - Before proceeding to a discussion of the perceived signal-to-noise ratio, we present an important theorem relating the visual capacity to a statistical quantity that is representative of the information content of actual observed scenes. We define the quantity G according to the formula,

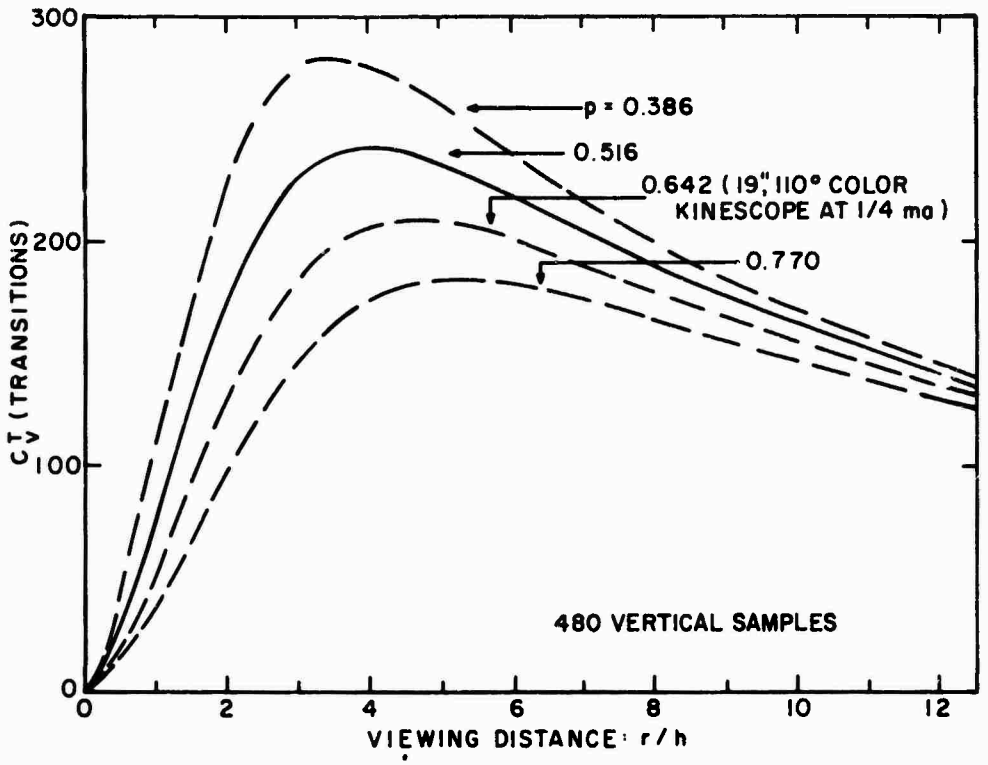


Figure 9. Visual capacity C_v^T as a function of viewing distance in picture heights for the vertical direction ($N_s = 480$) of a television display. The values of the parameter p, which represents the effective width of the Gaussian printing function, are given in the figure. The particular value obtained from measurements of a commercially available kinescope is indicated.

$$G = \frac{\langle (dE_s(x)/dx)^2 \rangle}{\langle I_o^2(x) \rangle} \tag{36}$$

Thus, G is the square of the gradient of the perceived signal response, averaged over the ensemble of scenes and normalized with respect to brightness. We call G the perceived *gradient content* of the ensemble of scenes. Its physical meaning can be grasped by considering the fact that visual information is contained only in regions where the perceived brightness varies with distance on the display screen (e.g., edges, boundaries, etc.). We can view G as the average inverse square of the distance required for a perceived transition. Thus, the gradient content may be regarded as a measure of the information density of the displayed scenes. The theorem we wish to prove is that, for the observed power spectrum of the ensemble of scenes $\Phi(\omega) \propto 1/\omega^2$,

$$G = (2\omega_L/w)\ C_v^T \tag{37}$$

where ω_L is an effective lower cutoff frequency of $\Phi(\omega)$. Equation (37) states that, at any viewing distance, the edge discrimination ability of the observer is proportional to the statistical average of a quantity that is representative of the perceived information content of the observed scenes. The importance of this result is that it relates C_v^T, which describes the system *capability*, to an actual response G that describes the average perceived *content* of the displayed scenes.

To prove Eq. (37), we first write the expression for $\langle(dE_S(x)/dx)^2\rangle$, using Eq. (16) for $E_S(x)$ and Eq. (23) for $R_{eff}(\omega)$

$$\begin{aligned}\langle(dE_S(x)/dx)^2\rangle = &\int_{-\omega_M}^{+\omega_M} \frac{d\omega}{2\pi} \int_{-\omega_M}^{+\omega_M} \frac{d\omega'}{2\pi}\ O(\omega r/2\pi)\ O(\omega' r/2\pi) \\ &\times R_{eff}(\omega)\ R_{eff}(\omega')\ \langle\hat{I}_o(\omega)\ \hat{I}_o(\omega')\rangle\ (i\omega)\ (i\omega') \\ &\times \exp\ [i(\omega + \omega')x]\end{aligned} \tag{38}$$

Next we employ Eq. (20) to relate the quantity within the bracket to the power spectrum $\Phi(\omega)$. This permits a trivial integration over ω'. We have then*

$$\langle(dE_S(x)/dx)^2\rangle = (1/\pi) \int_0^{\omega_M} d\omega\ O^2(\omega r/2\pi)\ |R_{eff}(\omega)|^2\ \omega^2\Phi(\omega) \tag{39}$$

In Section II we reported that the power spectrum of random scenes has the form

$$\Phi(\omega) = B/\omega^2; \quad |\omega| \gtrsim \omega_L \tag{40}$$

*We make use of the identity $R^*_{eff}(\omega) = R_{eff}(-\omega)$, a consequence of the reality of E(x), to change the integral over all ω to one over only positive frequencies.

Here ω_L, the effective lower cutoff frequency, corresponds to a spatial wavelength of roughly one picture width, so that $\omega_L \tilde{=} 2\pi/w$. The exact value of ω_L must be determined from the extreme low-frequency behavior of $\Phi(\omega)$, as discussed in Section II. If ω_L is much smaller than the frequency range of the integrand of Eq. (39)*, we can safely ignore the contribution to the integral from the frequency interval $|\omega| < \omega_L$. Combining Eqs. (39) and (40), and employing Eq. (27) for C_v^T, we have

$$\langle (dE_S(x)/dx)^2 \rangle = (B/w)\, C_v^T \tag{41}$$

We can express the power amplitude factor B in terms of the ensemble average input power $\langle I_o^2(x) \rangle$. The same procedure that was used above to calculate $\langle (dE_S(x)/dx)^2 \rangle$ can be employed to express $\langle I_o^2(x) \rangle$ in terms of $\Phi(\omega)$. We obtain

$$\langle I_o^2(x) \rangle = (1/\pi) \int_0^\infty d\omega\, \Phi(\omega) \tag{42a}$$

$$= (B/2\omega_L) \tag{42b}$$

where, in proceeding from Eq. (42a) to Eq. (42b), we have employed a simple Lorentzian, $\Phi(\omega)=B[\omega^2+\omega_L^2]^{-1}$, to describe the low-frequency behavior of $\Phi(\omega)$. Equations (40) and (42) may be regarded as constituting a precise definition of ω_L. Combining Eqs. (41) and (42b) immediately gives us the desired result, Eq. (37), with the gradient content G defined in Eq. (36).

Finally, we note that, since $\omega_L \tilde{=} 2\pi/w$, Eq. (37) can be carried one step further to give the approximate result

$$w^2\, G = 4\pi\, C_v^T \tag{43}$$

*With respect to $R_{eff}(\omega)$, this condition is automatically satisfied for $N_s >> 1$ and $N_{TV} >> 1$. This condition, when applied to $O(\omega r/2\pi)$, implies that $\omega_L^s r/2\pi << \nu_o$, or $r/w << 500$. Thus, we must confine ourselves to sufficiently small viewing distances for the peak of the eye's sensitivity curve to lie well above a frequency corresponding to one cycle/picture width. In practice this is the only interesting case.

The quantity on the left-hand side of Eq. (43) represents the mean square number of transitions perceived across the display by an observer at a particular viewing distance. This follows from the view of the quantity G as the average inverse square length required for a perceived transition. The quantity on the right-hand side represents, aside from the factor 4π, the number of perceivable edge transitions that the display-observer system is capable of producing. We emphasize the statistical relationship contained in Eq. (43) by rewriting the left-hand side in terms of N_{rms}, the statistically meaningful root mean square number of perceived edges associated with an arbitrary set of random natural scenes. From the above discussion, we have

$$N_{rms} = w\, G^{1/2} \tag{44}$$

in which case Eq. (43) becomes

$$N_{rms} = 2\,(\pi\, C_v^T)^{1/2} \tag{45}$$

The square root relationship between the root mean square number of perceived transitions and the maximum number that can be perceived is a direct consequence of the measured $1/\omega^2$ power spectrum.

2. The Perceived Signal-to-Noise Ratio

In Section III.C, we showed how the perceived response E(x) can be separated into the signal contribution $E_S(x)$, given by Eq. (16), and a sampling noise contribution $E_N(x)$, given by Eq. (17). It was proved that, after averaging over the ensemble of scenes and over position on the display screen, the product $E_N(x)\, I_o(x)$ vanishes, thereby showing that $E_N(x)$ is uncorrelated with the ensemble of input scenes. We can obtain a measure of the deleterious effect of the sampling noise by computing the signal-to-noise ratio at the perceptual level S/N. To do so we shall compute the mean square perceived signal "power" S^2 and the mean square perceived noise "power" N^2. Here, the term "power" refers to the square of the brightness; in this formulation, brightness is the analog of the current or the voltage of electrical communication theory.

We begin by writing the expression for the total perceived structural content $\overline{\langle E^2(x)\rangle}$. Using Eq. (13) to express $\overline{\langle E^2(x)\rangle}$ in terms of the contributions from the $E_m(x)$, defined in Eq. (14), we have

$$\overline{\langle E^2(x)\rangle} = \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} \overline{\langle E_m(x) E_n(x)\rangle} \tag{46}$$

By making use of the technique employed in proving the theorem of Section III.C, it is a simple matter to show that only the terms m = n contribute to the summations in Eq. (46). We readily find

$$\overline{\langle E^2(x)\rangle} = S^2 + N^2 \tag{47}$$

where*

$$S^2 = \overline{\langle E_S^2(x)\rangle} = \int_{-\omega_M}^{+\omega_M} \frac{d\omega}{2\pi}\, O^2(\omega r/2\pi)\, |\hat{\Pi}(\omega)|^2 \times \operatorname{sinc}^2(s\omega/\omega_s)\, |R(\omega)|^2\, \Phi(\omega) \tag{48}$$

$$N^2 = \overline{\langle E_N^2(x)\rangle} = \sum_{m\neq 0}^{\pm\infty} \int_{m\omega_s-\omega_M}^{m\omega_s+\omega_M} \frac{d\omega}{2\pi}\, O^2(\omega r/2\pi)\, |\hat{\Pi}(\omega)|^2 \times \operatorname{sinc}^2[s(\omega - m\omega_s)/\omega_s]\, |R(\omega - m\omega_s)|^2\, \Phi(\omega - m\omega_s) \tag{49}$$

We find it convenient to rewrite Eqs. (48) and (49), using Eq. (23) for the effective display MTF $R_{eff}(\omega)$. Making the dependence of S^2 and N^2 on viewing distance explicit, we have

*The equations for S^2 and N^2 do not contain factors arising from angular dispersion of the light emitted from the display screen. We assume that such factors are identical for both S^2 and N^2, so that they do not appear in the ratio S/N.

$$S^2(r) = \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi} \; O^2(\omega r/2\pi) \; |R_{eff}(\omega)|^2 \; \Phi(\omega) \tag{50}$$

$$N^2(r) = \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi} \; O^2(\omega r/2\pi) \; |R_{eff}(\omega)|^2 \; N(\omega) \tag{51}$$

Here $N(\omega)$ is the effective noise power spectrum, given by

$$N(\omega) = \sum_{m \neq 0}^{+\infty} \frac{\text{sinc}^2 \, [s(\omega - m\omega_s)/\omega_s] \; |R(\omega - m\omega_s)|^2}{\text{sinc}^2 \, (s\omega/\omega_s) \; |R(\omega)|^2} \; \Phi(\omega - m\omega_s) \tag{52}$$

Sampling Noise

Equations (50) and (51) are quite general. If, instead of sampling noise, we consider a noise power spectrum arising from other random noise sources associated with the transmission of the video signal, Eq. (51) still applies, if we use the proper form for $N(\omega)$. A physically interesting case is that of white noise [constant $N(\omega)$], for which we write

$$N(\omega) = \frac{N_s^2}{\int_{-\infty}^{+\infty} \frac{d\omega}{2\pi} \cdot |R_{eff}(\omega)|^2}; \quad \textit{White Noise} \tag{53}$$

Here N_s^2 is the mean square noise fluctuation, as measured on the display screen (obtained from Eq. (51) by setting $O(\omega r/2\pi) = 1$), and the quantity in the denominator is the noise equivalent bandwidth.

a. Properties of the Perceived Signal-to-Noise Ratio. - Equations (50) to (52) constitute the results for the signal-to-noise ratio for the one dimensional sampled display problem. To calculate the perceived signal-to-noise ratio, the integrals of Eqs. (50) and (51) must be evaluated and the ratio S/N formed. We note first that S/N is, in general, an explicit function of viewing distance. This arises from the fact that the noise power spectrum and the signal power spectrum are different functions of ω, so that the human visual system acts on these spectra differently at different viewing distances. Indeed, the

the measured $\Phi(\omega) = B/\omega^2$ implies that, if $N(\omega)$ falls off more slowly than $1/\omega^2$ at low frequencies, $S(r)/N(r)$ diverges at very large viewing distances. Although it is possible to conceive of pathological noise spectra that would give unusual behavior as r increases from zero, we find that $S(r)/N(r)$ normally increases monotonically with r. This is in agreement with common experience; when confronted with a noisy picture, we improve the perceived signal-to-noise ratio by moving away from the picture. By moving away, we use the MTF of the visual system to filter out the high-frequency noise components. This action is effective because, whereas the high-frequency part of the signal spectrum rolls off as $1/\omega^2$, the noise spectrum normally rolls off more slowly or even increases with frequency.

An example of the behavior of $S(r)/N(r)$ as a function of viewing distance is shown in Fig. 10 where the perceived signal-to-noise ratio is plotted for the vertical samples of the real television display described above in the discussion of the visual capacity and for which $C_v^T(r)$ is graphed in Fig. 8. It is seen from the figure that $S(r)/N(r)$ increases monotonically from a value near unity at $r = 0$. The sharper rise after the knee in the curve near $r/h = 1$ is due to the retinal frequency $\omega_s r/2\pi$ passing through the peak of the eye's sensitivity curve at ν_o. Since much of the noise spectrum is concentrated near $\omega = m\omega_s$ [see Eq. (52)], the perceived noise power falls off drastically for $\omega_s r/2\pi > \nu_o$, and $S(r)/N(r)$ increases rapidly. We find that, for large viewing distances, $r/h \gtrsim 8$, $S(r)/N(r)$ rises approximately as $(r/h)^2$. At the viewing distance $r_p/h = 4.7$ for which the visual capacity of the display has a maximum value (see Fig. 8), the value of S/N is approximately 21. We have used Eq. (53) for the power spectrum of white electronic noise to obtain an estimate of the perceived signal-to-noise ratio for the horizontal (analog) direction of the television display. This analysis indicates that, for high-quality (43 dB), commercial television pictures viewed at $r/h \tilde{=} 5$, S/N for the horizontal direction is also approximately 20.

From these results, it appears that the horizontal and vertical directions of the television display are well matched in terms of perceived signal-to-noise ratio; the sampling noise and analog picture noise are approximately equal. We also conclude that the value $S/N \tilde{=} 20$ is representative of what observers consider to be high-quality picture rendition. This value is consistent with the Weber-Fechner law, which indicates that intensity variations

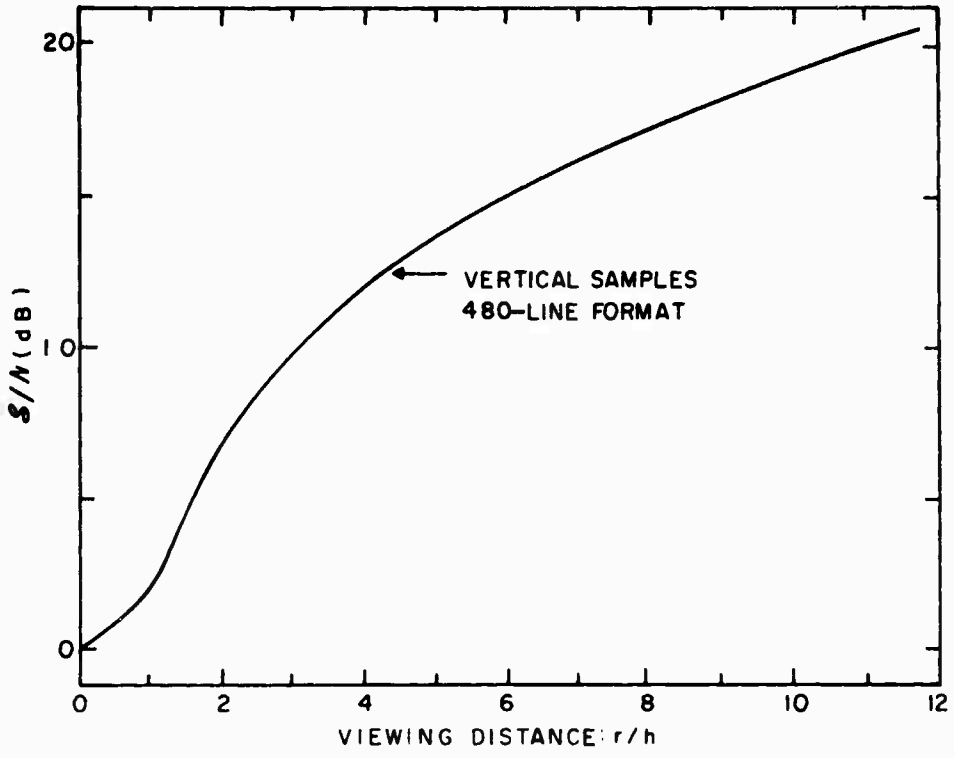


Figure 10. Perceived signal-to-noise ratio S/N as a function of viewing distance in picture heights for the vertical (sampled) direction of a real television display device. The C_V^T for this display is given in Fig. 8.

of approximately [14] 1% to 14% are perceivable, depending on the experimental conditions.

b. Asymptotic Expressions for S(r)/N(r). - Using Eqs. (50) to (52), we are able to obtain approximate analytic expressions for the asymptotic behavior of S(r)/N(r) at very large and very small viewing distances. Such expressions are very valuable because they help us to identify how S(r)/N(r) depends on the various display parameters as well as the viewing distance. In the following development, we illustrate the technique employed to calculate asymptotic expressions for S(r)/N(r) by considering a special case corresponding to a constant printing function that extends over the entire sampling location; that is,

$$P(x) = 1; \text{ for } |x| \leq \frac{1}{2} x_o$$

$$= 0; \text{ for } |x| > \frac{1}{2} x_o$$

$$\hat{\Pi}(\omega) = \text{sinc } (\omega/\omega_s) \tag{54}$$

We shall then present, without derivation, the results for Gaussian printing functions. The case $\hat{\Pi}(\omega) = \text{sinc}\ (\omega/\omega_s)$ is of special interest because it represents the simplest printing function for the optical block processor described elsewhere in this report. It has also been used extensively in the computer generation [4] of sampled images.

In calculating the asymptotic behavior of $S(r)/N(r)$, we consider viewing distances much larger than the distance r_p for maximum C_v^T, given approximately by Eq. (34). However, we shall restrict ourselves to sufficiently small viewing distances in order for the peak of the eye's sensitivity curve at $\nu = \nu_o$ to lie well above the retinal frequency $\omega_L r/2\pi \tilde{=} r/w$. Under these conditions, in Eq. (50) for $S^2(r)$ we can replace $R_{eff}(\omega)$ by unity and substitute for $\Phi(\omega)$ the measured signal power spectrum B/ω^2. After transforming the integration variable to the retinal frequency coordinate $\nu = \omega r/2\pi$, Eq. (50) becomes

$$S^2(r) = (Br/2\pi^2) \int_0^\infty d\nu\ \nu^{-2}\ O^2(\nu);$$

$$\text{for } \Phi(\omega) = B/\omega^2,\ r_p << r << \nu_o w \qquad (55)$$

The integral is finite since $O(\nu) \propto \nu$ as $\nu \to 0$ [see Fig. 5 and Eq. (32)]. Equation (55) is remarkable in itself, for it shows that, aside from factors arising from angular dispersion of the light emitted from the display screen, the perceived signal power becomes larger as we increase the viewing distance. This result is a simple consequence of the concentration of the signal power of natural scenes at low frequencies. As r increases, the peak of the MTF of the visual system at $\nu = \nu_o$ corresponds to lower and lower frequencies as measured on the display screen, thereby permitting more efficient overall transmission of the signal power. We note that if the power spectrum of natural scenes were white, Eq. (50) would give us

$$\lim_{r \to \infty} S^2(r) = C/\theta_e^\infty r;\ \text{for } \Phi(\omega) = C \qquad (56)$$

Here θ_e^∞ is the perceived angular width of an edge transition, given by Eq. (29). In this case, the perceived signal decreases with r.

Turning now to Eqs. (51) and (52) for the noise power $N^2(r)$, substitution of $\Phi(\omega) = B/\omega^2$ and using Eq. (54) for $\hat{\Pi}(\omega)$ and Eq. (23) for $R_{eff}(\omega)$ gives us

$$N^2(r) = B \sum_{m\neq 0}^{\pm\infty} \int_{m\omega_s-\omega_M}^{m\omega_s+\omega_M} \frac{d\omega}{2\pi} O^2(\omega r/2\pi) \left|R(\omega-m\omega_s)\right|^2 \frac{\text{sinc}^2(\omega/\omega_s)}{(\omega - m\omega_s)^2}$$

$$\times \text{sinc}^2[s(\omega - m\omega_s)/\omega_s] \qquad (57)$$

In Eq. (57), it was not necessary to employ the lower cutoff frequency ω_L of the $1/\omega^2$ power spectrum. This is because the quantity $[\text{sinc}^2(\omega/\omega_s)]/(\omega-m\omega_s)^2$ is finite at $\omega = m\omega_s$; no singularity problem arises from the extension of the argument of $\Phi(\omega)$ to zero. This procedure gives negligible error, since ω_L is much less than the range of the various functions in the integrand of Eq. (57). However, this procedure is possible only for the particular functions $\hat{\Pi}(\omega) = \text{sinc}(n\omega/\omega_s)$ with n = integer, since they are the only functions [10] that (1) are the Fourier transforms of real positive functions and (2) have zeroes at $\omega = m\omega_s$. For all other functions $\hat{\Pi}(\omega)$, the lower cutoff frequency must be explicitly employed to prevent singular behavior at $\omega = m\omega_s$. The physical basis for this result lies in the fact that the printing functions corresponding to $\hat{\Pi}(\omega) = \text{sinc}\,(n\omega/\omega_s)$ completely and uniformly illuminate n sampling locations. This confines the perceived noise to the vicinity of the edge transitions that give rise to the $1/\omega^2$ power spectrum. The average noise is then independent of the display width and, therefore, also independent of ω_L. All other printing functions produce "ripple" fluctuations across the display screen, so that the perceived noise power arising from these fluctuations depends explicitly on ω_L.

We can simplify Eq. (57) by changing the sum over all integers m to one over only positive integers. This is readily accomplished by using the fact that each of the functions appearing in the integrand are even with respect to a sign change of their arguments. Next, for the sake of simplicity, we use $R(\omega) = 1$. Then, Eq. (57) becomes

$$N^2(r) = (B/\pi) \sum_{m=1}^{\infty} \int_{m\omega_s-\omega_M}^{m\omega_s+\omega_M} d\omega\, O^2(\omega r/2\pi) \frac{\text{sinc}^2(\omega/\omega_s)}{(\omega-m\omega_s)^2} \text{sinc}^2[s(\omega-m\omega_s)/\omega_s] \qquad (58)$$

In general, Eq. (58) must be evaluated numerically. However, for displays operating in the extreme oversampled or in the undersampled limit, we can obtain approximate expressions for $N^2(r)$.

In the extreme oversampled limit $\omega_s >> \omega_M$, it is seen from Eq. (58) that the noise spectrum consists of the sum of narrow-band noise contributions, each of width $2\omega_M$, centered about multiples of the sampling frequency $m\omega_s$. In that case, the integrand of Eq. (58) may be replaced by its value at $\omega = m\omega_s$. The integration is then trivial, and we are left with

$$N^2(r) = (2B\omega_M/\pi\omega_s^2) \sum_{m=1}^{\infty} O^2(m\omega_s r/2\pi)/m^2; \text{ for } \omega_M << \omega_s \tag{59}$$

Next, for large viewing distances such that $r/2\pi\nu_o >> 1/\omega_s$, we can replace the MTF of the human visual system by its high-frequency, inverse-square rolloff characteristic [11]

$$O(\nu) = 1.42\ (\nu_o/\nu)^2; \text{ for } \nu >> \nu_o \tag{60}$$

Employing Eq. (60) in Eq. (59), and making use of the mathematical identity [15] $\sum_{m=1}^{\infty} (1/m^6) = \pi^6/945$, we obtain our result for the perceived noise power at large viewing distances for the oversampled limit

$$N^2(r) = 1.31\ B\ (\omega_M/\omega_s^2)\ (2\pi\nu_o/\omega_s r)^4;$$
$$\text{for } \Phi(\omega) = B/\omega^2,\ \omega_M << \omega_s,\ r_p << r << \nu_o w \tag{61}$$

Using Eqs. (55) and (61) for $S^2(r)$ and $N^2(r)$, respectively, we obtain, for the perceived signal-to-noise ratio,

$$S(r)/N(r) = 3.38 \times 10^{-7}\ (N_s^3/N_{TV}^{1/2})\ (r/w)^{5/2};$$
$$\text{for } \Phi(\omega) = B/\omega^2,\ N_{TV} << N_s.\ r_p << \nu_o\ w \tag{62}$$

To obtain Eq. (62), we have employed the value calculated from the measured $O(\nu)$ (Fig. 5),

$$\int_0^\infty d\nu\ \nu^{-2}\ O^2(\nu) = 4.73/\nu_o \tag{63}$$

and have used Eqs. (2) and (35) to express ω_s and ω_M in terms of the number of samples N_s and number of TV lines N_{TV}, respectively.

In the undersampled limit $\omega_s < \omega_M$, the lower limits of the integrals of Eq. (58) extend down to dc provided $m \leq \text{Mod}\ (\omega_M/\omega_s)$, where Mod(x) stands for the "greatest integer in x". Then, for sufficiently large viewing distances such that the inequality $r/2\pi\nu_o >> 1/\omega_s$ is satisfied, the major contribution to $N^2(r)$ arises from m-values up to and including $m = \text{Mod}(\omega_M/\omega_s)$. The contributions for $m > \text{Mod}(\omega_M/\omega_s)$ are assumed to lie sufficiently far above the peak of the eye's sensitivity curve at $\nu = \nu_o$ for them to be neglected. Of course, the discontinuity in allowed m-values when ω_M/ω_s = integer is not strictly correct. In practice a continuous transition occurs due to the fact that the viewing distance is not formally infinite. The situation for the undersampled display is to be compared with that for the oversampled display, where the only noise contributions were centered in narrow bands around $\omega = m\omega_s$, and, therefore, are perceived with far less subjective intensity at large viewing distances. In the undersampled case, the ability of the display to fold high frequencies down to dc generates, on the average, perceivable noise well within the effective passband of the human visual system even at large viewing distances. With these considerations, we approximate Eq. (58) for $N^2(r)$ as follows:

$$N^2(r) = (B/\pi\omega_s^2) \int_{-\infty}^{+\infty} d\omega\ O^2(\omega r/2\pi)\ F_s(\omega,\ \omega_M/\omega_s);$$

$$\text{for } \omega_M > \omega_s,\ r_p << r << \nu_o\ w \tag{64}$$

where

$$F_s(\omega,\omega_M/\omega_s) = (\omega/\omega_s)^2 \sum_{m=1}^{\text{Mod}(\omega_M/\omega_s)} 1/m^4; \text{ for } s = \text{integer} \qquad (65a)$$

$$= \sum_{m=1}^{\text{Mod}(\omega_M/\omega_s)} [\text{sinc}^2(ms)]/m^2; \text{ for } s \neq \text{integer} \qquad (65b)$$

In arriving at Eqs. (64) and (65) from Eq. (58), we have: (1) replaced the term $(\omega-m\omega_s)^2$ by $(m\omega_s)^2$, (2) kept only the first non-vanishing term in the expansion of sinc $[s(\omega-m\omega_s)/\omega_s]$ about $\omega = 0$, (3) replaced sinc (ω/ω_s) by unity, and (4) extended the integration limits* to $\pm\infty$. All these approximations are valid for $r >> r_p$.

We first treat the case s = integer. The sum $F_s(\omega, \omega_M/\omega_s)$ in Eq. (65a) depends only weakly on ω_M/ω_s, varying by 8% over the range $\omega_s < \omega_M < \infty$. Therefore, we shall use the $\omega_M/\omega_s = \infty$ limit [15]

$$F_s(\omega,\infty) = (\pi^4\omega^2/90\omega_s^2); \text{ for } s = \text{integer} \qquad (66)$$

in the calculation of $N^2(r)$. Inserting Eq. (66) into Eq. (64) and transforming the frequency variable of integration to the retinal frequency coordinate $\nu = \omega r/2\pi$, we obtain

$$N^2(r) = (\pi^3 B/45\omega_s)\ (2\pi/\omega_s r)^3 \int_0^\infty d\nu\ \nu^2\ O^2(\nu);$$

$$\text{for } \Phi(\omega) = B/\omega^2, \omega_M > \omega_s,\ r_p << r << \nu_o w,\ s = \text{integer} \qquad (67)$$

*By so doing, we implicitly exclude the particular case $\omega_s = \omega_m$, for which we should integrate over only positive frequencies.

Now, using Eq. (55) for $S^2(r)$, we have

$$S^2(r)/N^2(r) = \frac{45}{\pi 4} [N_s(r/w)]^4 \frac{\int_0^\infty d\nu\ \nu^{-2}\ O^2(\nu)}{\int_0^\infty d\nu\ \nu^2\ O^2(\nu)};$$

$$\text{for } \Phi(\omega) = B/\omega^2,\ N_{TV} > 2N_s,\ r_p << r << \nu_o w,\ s = \text{integer} \quad (68)$$

Here we have again employed Eqs. (2) and (35) to express ω_s and ω_M in terms of N_s and N_{TV}, respectively. Utilizing the numerical result,

$$\int_0^\infty d\nu\ \nu^2\ O^2(\nu) = 4.53\ \nu_o^3 \quad (69)$$

obtained from the measured $O(\nu)$ (Fig. 5) along with Eq. (63) for the quantity $\int_0^\infty d\nu\ \nu^{-2}\ O^2(\nu)$, Eq. (68) reduces to

$$S(r)/N(r) = 3.31 \times 10^{-6} [N_s(r/w)]^2;$$

$$\text{for } \Phi(\omega) = B/\omega^2,\ N_{TV} > 2N_s,\ r_p << r << \nu_o w,\ s = \text{integer} \quad (70)$$

For the case $s \neq$ integer, we again find that the sum $F_s(\omega, \omega_M/\omega_s)$, defined in Eq. (65b), varies relatively slowly with ω_M/ω_s. For $s = 0$, $F_s(\omega, \omega_M/\omega_s)$ rises monotonically from unity to $\pi^2/6$ over the range $\omega_s < \omega_M < \infty$. Similarly, for values of s near unity, we find that the ratio $F_s(\omega,\infty)/F_s(\omega,i)$ has the value $\pi^2/6$. At intermediate values of s, the relative variation of $F_s(\omega,\omega_M/\omega_s)$ with ω_M/ω_s is somewhat smaller. For purposes of illustration, we will take the limit [16],

$$F_s(\omega,\infty) = \pi^2(1 - s)^2/6;\ \text{for } s < 1 \quad (71)$$

with the realization that we are ignoring a weak variation with ω_M/ω_s that, in the worst case, will give an error of less than 30% in $S(r)/N(r)$. Substituting Eq. (71) into Eq. (64) for $N^2(r)$ readily gives the result

$$N^2(r) = \pi^2 B(1 - s)^2/3\omega_s{}^2\, \theta_e^{\infty} r;$$

$$\text{for } \Phi(\omega) = B/\omega^2,\ N_{TV} > 2\ N_s,\ r_p << r << \nu_o w,\ s < 1 \qquad (72)$$

The perceived signal-to-noise ratio in the case at hand is obtained from Eq. (72) along with Eq. (55) for $S^2(r)$. Utilizing Eq. (63) and the value of θ_e^{∞} from Eq. (30), we find

$$S(r)/N(r) = 1.83 \times 10^{-3}\ N_s(r/w)/(1 - s);$$

$$\text{for } \Phi(\omega) = B/\omega^2,\ N_{TV} > 2\ N_s,\ r_p << r << \nu_o w,\ s < 1 \qquad (73)$$

Equations (62), (70), and (73) constitute the essential results of the perceived signal-to-noise problem at large viewing distances for $\hat{\Pi}(\omega) = \text{sinc}\ (\omega/\omega_s)$. In discussing these results, we notice first that all the expressions for $S(r)/N(r)$ diverge at large viewing distances. This is true even for the undersampled displays which contribute noise power at the lowest frequencies. Although this result may have been expected on intuitive grounds, it is indeed nontrivial, since it is a direct consequence of the measured $1/\omega^2$ power spectrum of natural scenes. To emphasize this point in a dramatic way, we shall calculate $S(r)/N(r)$ for undersampled displays using a white signal power spectrum and show that a catastrophe results. The calculation is easily performed by going back to Eq. (58) for the general $N^2(r)$ with $R(\omega) = 1$ and $\hat{\Pi}(\omega) = \text{sinc}\ (\omega/\omega_s)$. We replace the quantity $B/(\omega - m\omega_s)^2$, representing the power spectrum, by a constant C. Let us consider the special case $s = 0$ and take delta-function printing. These statements require that we replace the two sinc functions in Eq. (58) by unity. With these modifications, the formula for $N^2(r)$ becomes

$$N^2(r) = (C/\pi) \sum_{m=1}^{\infty} \int_{m\omega_s - \omega_M}^{m\omega_s + \omega_M} d\omega\ O^2(\omega r/2\pi);$$

$$\text{for } \Phi(\omega) = C,\ s = p = 0 \qquad (74)$$

Now, for an undersampled display $\omega_s < \omega_M$ and for $r >> r_p$, we can cut off the sum at $m = \text{Mod}\ (\omega_M/\omega_s)$ and extend the limits of integration to $\pm\infty$, as was done for the case of the $1/\omega^2$ power spectrum. In that case, Eq. (74) reduces to

$$\lim_{r \to \infty} N^2(r) = (2C/\theta^{\infty} r)\ \text{Mod}\ (\omega_M/\omega_s);$$

$$\text{for } \Phi(\omega) = C,\ N_{TV} > 2\ N_s,\ s = p = 0 \qquad (75)$$

Combining the above result with Eq. (56) for $S^2(r)$ for a white power spectrum, we have

$$\lim_{r \to \infty} S(r)/N(r) = [2\ \text{Mod}\ (\omega_M/\omega_s)]^{-1/2}$$

$$\text{for } \Phi(\omega) = C,\ N_{TV} > 2N_s,\ s = p = 0 \qquad (76)$$

Equation (76) predicts that the limiting perceived signal-to-noise ratio for an undersampled display with $s = p = 0$ is less than unity. Indeed, in the extreme undersampled limit, there is essentially "infinite Moiré power," and the asymptotic value of $S(r)/N(r)$ vanishes! This result is model-independent since it makes no statement about the observer except that he has an acuity limit. If the result were true, the simple act of placing a screen with extremely small holes* over a natural scene would produce zero perceived signal-to-noise ratio no matter how far away from the screen the observer positioned himself. Yet we know this will not happen, and the reason, simply put, is that the statistical property of natural scenes, as represented by the actual $1/\omega^2$ power spectrum, prevents it.

Returning now to a discussion of the results for $S(r)/N(r)$, Eqs. (62), (70), and (73), we can make several observations. First, all the $S(r)/N(r)$ diverge at large viewing distances as a power of r/w; the value of the exponent depends on whether the display is oversampled or undersampled and on the value

*The holes are presumed to be sufficiently large for diffraction effects to be negligible.

of the sampling width s. The perceived signal-to-noise ratio diverges most quickly [as $(r/w)^{5/2}$] for oversampled displays and least quickly [as (r/w)] for undersampled displays with $s < 1$. Second, the dependence of $S(r)/N(r)$ on the number of samples is also a power law, ranging from N_s^3 for oversampled displays to N_s for undersampled displays with $s < 1$. Third, at large viewing distances $S(r)/N(r)$ is independent of s for extremely oversampled displays but varies as $(1-s)^{-1}$ for undersampled displays with $s < 1$. As $s \to 1$, Eq. (73) no longer applies, and Eq. (70) must be employed. Thus, $S(r)/N(r)$ is sensitive to s only for the case of undersampled displays, where large or integer values of s are preferred. This behavior is to be compared with the visual capacity which stresses small values of s for increased edge appreciation ability.

Table 1 summarizes our results for $S(r)/N(r)$ in the large viewing distance limit for the constant printing function of Eq. (54) and for a Gaussian printing function

$$P(x) = \exp\left[-(x/px_o)^2\right]$$
$$\hat{\Pi}(\omega) = \exp\left[-(\pi p\omega/\omega_s)^2\right] \tag{77}$$

Here p, the effective printing width, is the distance in units of x_o required for $P(x)$ to fall to the value $1/e$. The expressions for $S(r)/N(r)$ for the Gaussian printing function were obtained using the technique employed in the derivations of the results for the constant printing function Eq. (54). The value $\omega_L = 2\pi/w$ was employed throughout, and a Lorentzian $\Phi(\omega) = B[\omega^2+\omega_L^2]^{-1}$ was used to describe the low-frequency behavior of the power spectrum. It is seen from Table 1 that, for oversampled displays, $S(r)/N(r)$ again diverges as $(r/w)^{5/2}$, as in the case of the constant printing function. Also, $S(r)/N(r)$ is again independent of s. However, the dependence on p is very rapid, increasing as $\exp(\pi^2p^2)$. Thus, large values of the printing width are preferred here. Notice that, for a given value of N_s and N_{TV}, there exists a critical value of p above which the Gaussian printing function gives a superior $S(r)/N(r)$ over the constant printing function Eq. (54). Denoting this critical value by p_c, we have from Table 1

$$\pi^2 p_c^2 = \ln\left(1.76\, N_s/N_{TV}^{1/2}\right); \text{ for } N_s >> N_{TV} \tag{78}$$

Table 1. Limiting Behavior of $S(r)/N(r)$ for $r \gg r_p$

Display Type	s	$\hat{\Pi}(\omega)$	$S(r)/N(r)$
Oversampled:	any value	sinc (ω/ω_s)	$3.38 \times 10^{-7}\ (N_s^3/N_{TV}^{1/2})\ (r/w)^{5/2}$
$(\omega_s \gg \omega_M)$	any value	$\exp\ [-\ (\pi p\omega/\omega_s)^2]$	$1.92 \times 10^{-7} N_s^2\ [\exp\ (\pi^2 p^2)]\ (r/w)^{5/2}$*
Undersampled:	integer	sinc (ω/ω_s)	$3.31 \times 10^{-6}\ [N_s(r/w)]^2$
$(\omega_s < \omega_M)$	$s < 1$	sinc (ω/ω_s)	$1.83 \times 10^{-3}\ N_s(r/w)/(1-s)$
	integer	$\exp\ [-\ (\pi p\omega/\omega_s)^2]$	$3.31 \times 10^{-6}\ [N_s(r/w)]^2$**
	$s < 1$	$\exp\ [-\ (\pi p\omega/\omega_s)^2]$	$1.83 \times 10^{-3}\ N_s(r/w)/(1-s)$†

*For non-Gaussian printing functions, replace the quantity $\exp[\pi^2 p^2]$ by $[\hat{\Pi}(\omega_s)]^{-1}$ provided $\hat{\Pi}(\omega)$ falls off sufficiently rapidly for the inequality $\hat{\Pi}2(2\omega_s)/16 \ll \hat{\Pi}^2(\omega_s)$ to be obeyed.

**For $r/w \gg 300 \exp\ [-2\pi^2 p^2]$; otherwise use $S(r)/N(r) = 1.92 \times 10^{-7}\ N_s^2\ [\exp\ (\pi^2 p^2)]\ (r/w)^{5/2}$. (Note: for $2\pi^2 p^2 \ll 1$, $r/w \gg 300$ is not allowed, since it violates the condition $r/w \ll \nu_o = 458$ assumed during the calculation.)

†For $r/w \gg 450\ [N_s(1-s) \exp\ (\pi^2 p^2)]^{-2/3}$; otherwise use $S(r)/N(r) = 1.92 \times 10^{-7}\ N_s^2\ [\exp\ (\pi^2 p^2)]\ (r/w)^{5/2}$.

As an example, we take $N_s = 4N_{TV} = 1000$, in which case Eq. (78) gives $p_c = 0.69$. For values of p above p_c the Gaussian printing function is sufficiently broad for the "ripple" fluctuations across the display screen produced by the Gaussian to be reduced to the point where the overall perceived noise power is smaller than that for the constant printing function. For the case of undersampled displays, close examination of Table 1 shows that, for s = integer, as p is increased, $S(r)/N(r)$ first follows the same law as for oversampled displays but eventually saturates at a value given by the $S(r)/N(r)$ for undersampled displays with a constant printing function. Undersampled displays with $s < 1$ are likely to follow the same law as for the constant printing function (independent of p), depending on the specific values of N_s, s, and r/w. The reason for the seemingly complicated behavior for undersampled displays lies in the competition between the noise spectra centered around multiples of the sampling frequency and the noise spectra concentrated at lower frequencies $\omega \lesssim 2\pi\ \nu_o/r \ll \omega_s$.

The former noise source has a much larger power spectrum $N(\omega)$, but the latter is transmitted to the perceptual level more efficiently because it lies well within the effective passband of the human visual system. As p is increased, we essentially filter out the high-frequency ($m\omega_s$) noise by decreasing the range of $\hat{\Pi}(\omega)$. Eventually, the low-frequency noise dominates, giving us the same results as for the case $\hat{\Pi}(\omega) = \text{sinc}\ (\omega/\omega_s)$.

The behavior of $S(r)/N(r)$ in the limit $r \to 0$ is of little practical interest. Furthermore, in actual fact, the calculation breaks down as $r \to 0$ because the MTF of the human visual system degrades as a result of the eye's inability to focus. Nevertheless, trends exhibited by $S(r)/N(r)$ as a function of the display parameters are important, so that we have employed Eqs. (50) to (52) to compute $S(0)/N(0)$. The results are presented without derivation in Table 2 for the cases of constant [Eq. (54)] and Gaussian [Eq. (77)] printing functions. We note first the apparently paradoxical inverse relationship between $S(0)/N(0)$ and N_s for the case of the Gaussian printing function with $\exp(2\pi^2 p^2) >> 1$. This result requires special comment. It does not violate the requirement that $S(r)/N(r) \to \infty$ as $N_s \to \infty$ since the entries in the table were calculated in the limit $\omega_s r/2\pi = N_s r/w \to 0$. Taking $N_s \to \infty$ before specifying r always gives $S(r)/N(r) \to \infty$. The inverse dependence of $S(0)/N(0)$ on N_s arises from the fact that, for $\omega_s r/2\pi \to 0$, the frequency ω_s corresponds to a retinal frequency $\omega_s r/2\pi$ that lies on the rising part of the MTF of the human visual system (see Fig. 5). Since much of the noise spectrum for the Gaussian printing function is concentrated near ω_s, the perceived noise power increases with the sampling rate as ω_s^2 because of the linear low-frequency behavior of $O(\nu)$ [Eq. (32)]. On the other hand, the effective passband of the perceived signal power is independent of ω_s for the oversampled case and is proportional to ω_s to only the first power in the undersampled case. Thus, the linear low-frequency behavior of $O(\nu)$ is responsible for the inverse dependence of $S(0)/N(0)$ on N_s.

Further examination of Table 2 shows: (1) for delta-function printing $p = 0$, $S(r)/N(r) \to 0$ as $r \to 0$, the exact dependence on r depending on the details of the sampling process, (2) for full-width printing $\hat{\Pi}(\omega) = \text{sinc}\ (\omega/\omega_s)$, $S(r)/N(r) \to 0$ as $r^{1/2}$, (3) $S(0)/N(0)$ increases rapidly with the printing width p, and (4) $S(0)/N(0)$ is sensitive to s only in the case of undersampled displays with large p, where large or integer values are preferred.

Table 2. Limiting Behavior of $S(r)/N(r)$ for $r \to 0$

Display Type	s	$\hat{\Pi}(\omega)$	S(r)/N(r)
Oversampled:	any value	$\text{sinc}\ (\omega/\omega_s)$	$S(r)/N(r) \to 0$ as $r^{1/2}$
($\omega_s >> \omega_M$)	any value	$\exp\ [-(\pi p\omega/\omega_s)^2]$	$[\exp\ (\pi^2p^2)]\ (N_{TV}/2\pi N_s^2)^{1/2}$ for $\exp\ (2\pi^2p^2) >> 1$
			$S(r)/N(r) \to 0$ as $r^{3/2}$ for $p = 0$
Undersampled:	any value	$\text{sinc}\ (\omega/\omega_s)$	$S(r)N(r) \to 0$ as $r^{1/2}$
($\omega_s << \omega_M$)	s = 0	$\exp\ [-(\pi p\omega/\omega_s)^2]$	$S(r)N(r) \to 0$ as r for $p = 0$
	s > 0	$\exp\ [-(\pi p\omega/\omega_s)^2]$	$S(r)N(r) \to 0$ as $r^{3/2}$ for $p = 0$
	any value	$\exp\ [-(\pi p\omega/\omega_s)^2]$	$\exp\ (\pi^2p^2)/[(2\pi)^{3/4}\ (N_sp)^{1/2}]$*

*For $1 << \exp\ (2\pi^2p^2) << 3800\ p^5N_s$ for s = integer or $200p^3\ N_s/(1-s)^2$ for $s < 1$; if the last part of the inequality is not satisfied, use $S(0)/N(0) = 2(3)^{1/2}\ p/(1-s)$ for $s < 1$ or $S(0)/N(0) = 4(15)^{1/2}p^2$ for s = integer.

c. Discussion Summary. - To summarize briefly the results of our analysis of the perceived signal-to-noise ratio S(r)/N(r), we have found that S(r)/N(r) increases monotonically with viewing distance. We have also found that S(r)/N(r) increases rapidly with the number of samples at ordinary viewing distances. Wide printing functions are definitely preferred, as are large or integer values of the sampling width s. However, the parameter s is important in determining S(r)/N(r) only in the case of undersampled displays. The dependence of S(r)/N(r) on s and p is opposite to that of the visual capacity, where small values of s and p are desired. Thus, we have the general rule: *Narrow width sampling and printing enhances the appreciation of sharpness through the increase of* $C_v^T(r)$. *However, it does so at the expense of the perceived signal-to-noise ratio* S(r)/N(r).

d. Calculated Examples. - Some of the important properties of the perceived signal-to-noise ratio are illustrated in Figs. 11 through 14, where we have employed Eqs. (50) to (52) to calculate S/N numerically for various display systems. The curves in Fig. 11 are the calculated S(r)/N(r) vs viewing distance in units of picture height h for a hypothetical television display. In

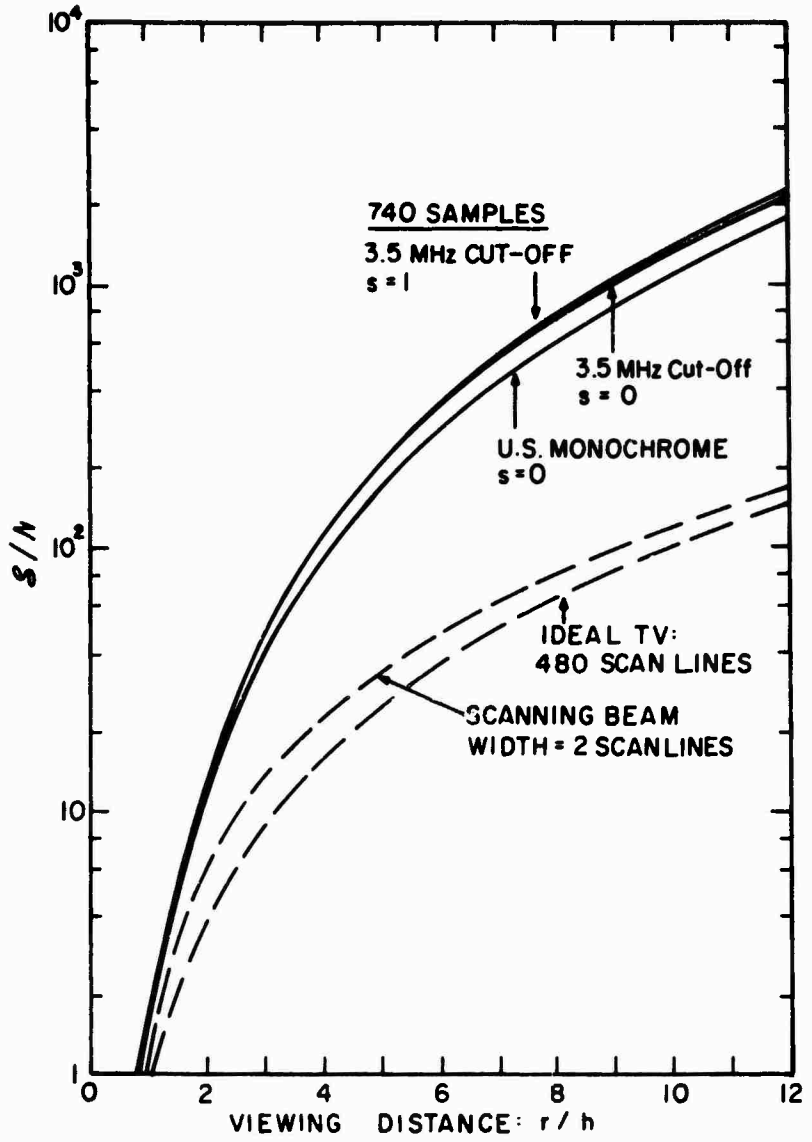


Figure 11. Perceived signal-to-noise ratio S/N as a function of viewing distance in picture heights for a hypothetical television display. The dashed curves represent the S/N for the vertical direction (N_s = 480) with two different printing functions (see text). The solid curves represent the S/N for the horizontal direction (N_s = 740) with two types of bandwidth limitation (see text). The values of the sampling width s are given in the figure.

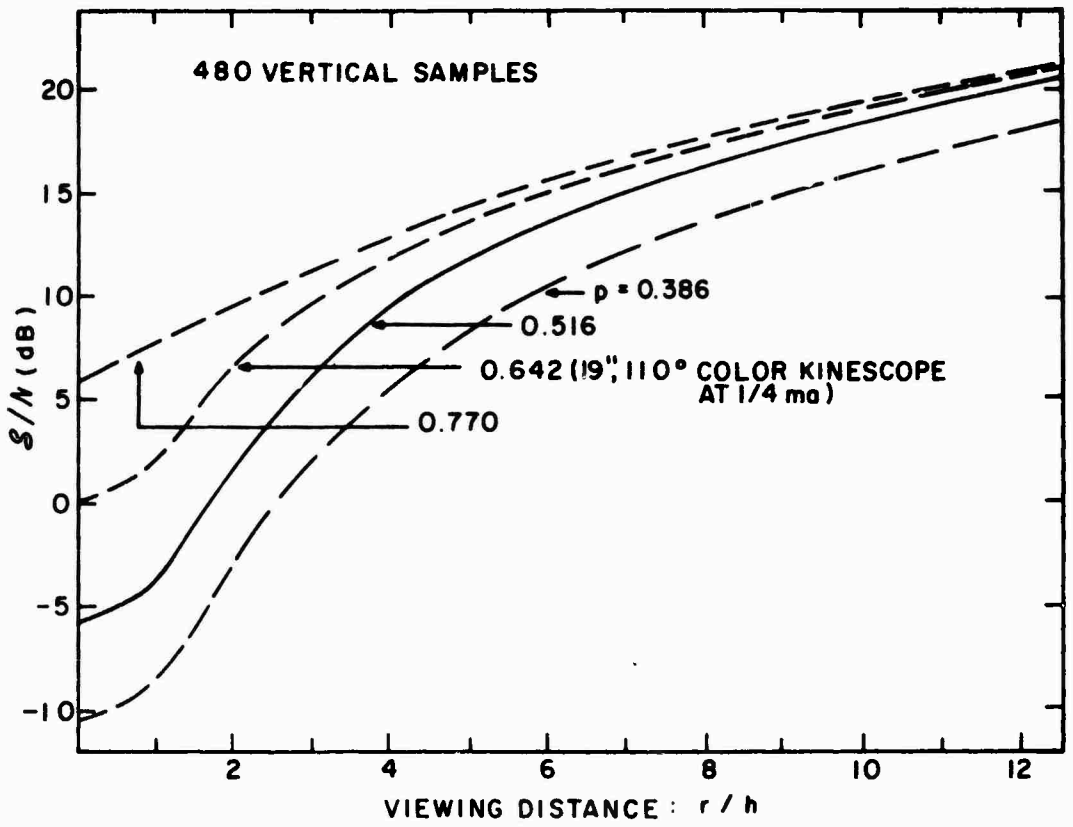


Figure 12. Perceived signal-to-noise ratio S/N as a function of viewing distance in picture heights for the vertical direction (N_s = 480) of a television display. The values of the parameter p, which represents the effective width of the Gaussian printing function, are given in the figure. The particular value obtained from measurements of a commercially available kinescope is indicated.

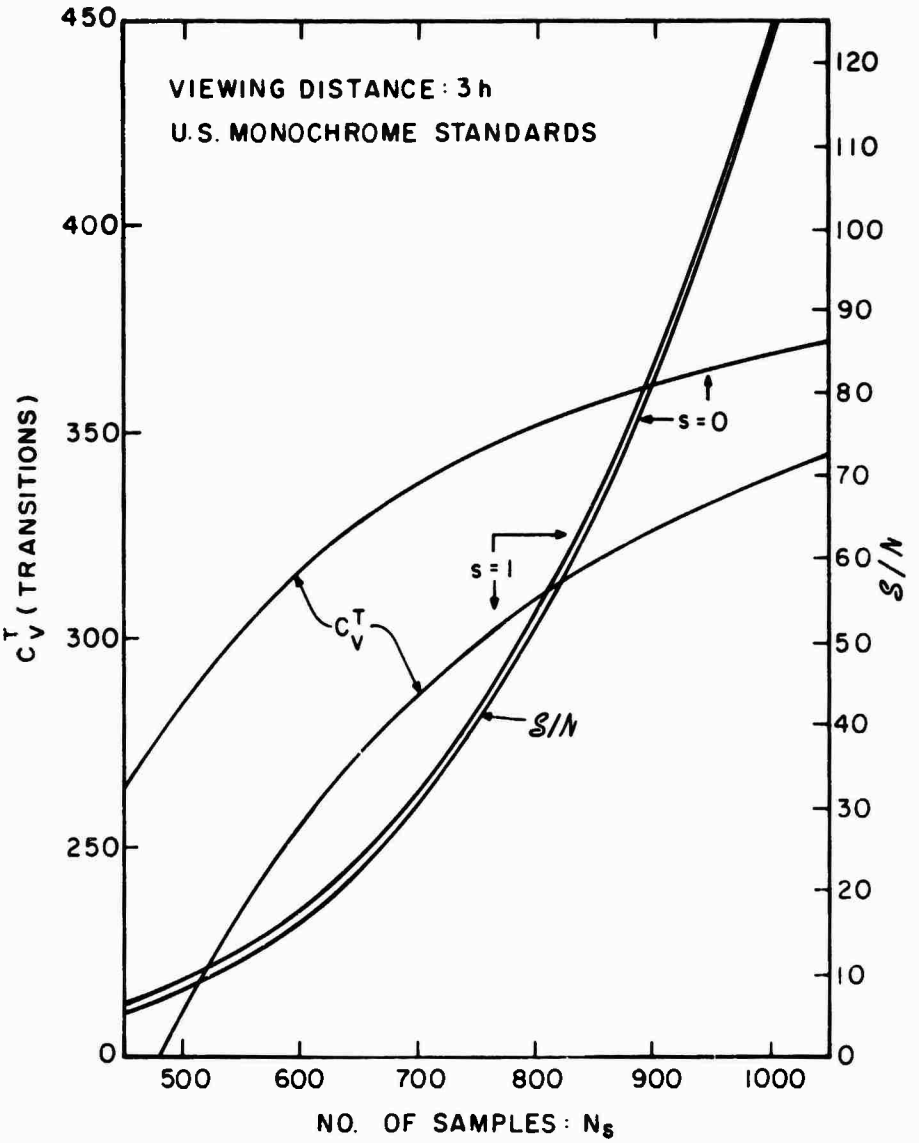


Figure 13. Visual capacity C_V^T and perceived signal-to-noise ratio S/N as a function of the number of samples for the horizontal direction of a hypothetical television display. The viewing distance is 3 picture heights, and the bandwidth limitation has been set by the U.S. Monochrome Standards (N_{TV} = 471). Curves are shown for s = 0 and s = 1.

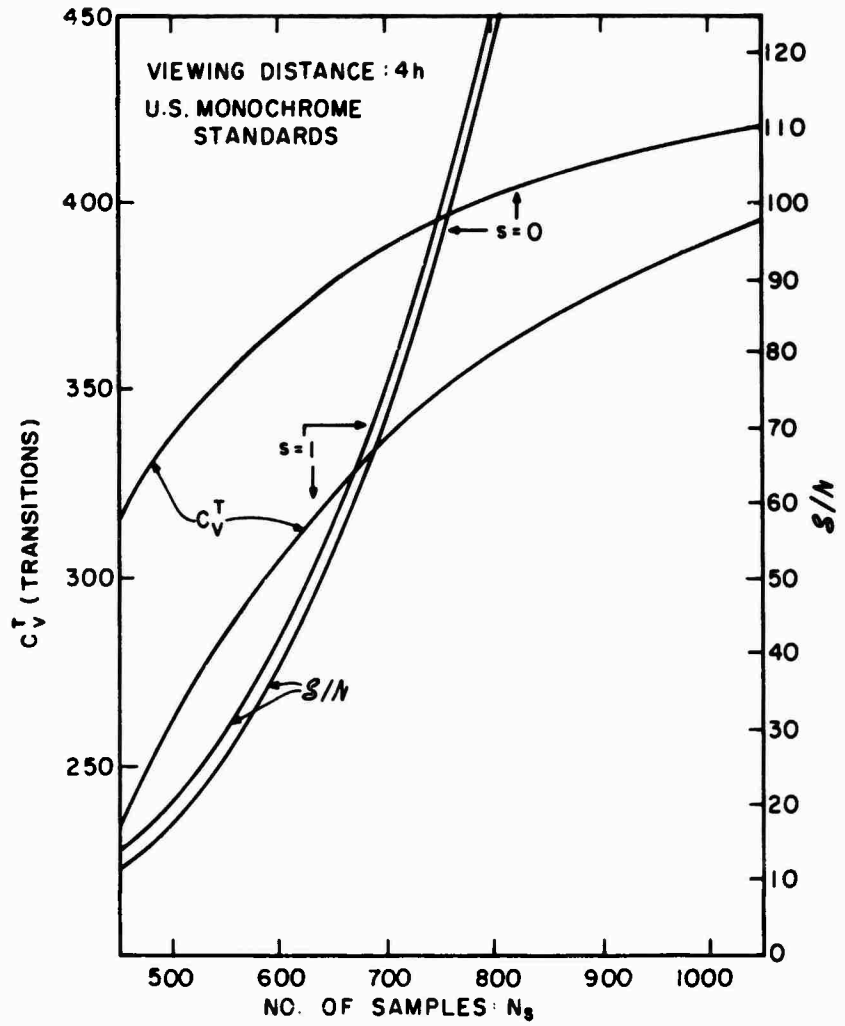


Figure 14. Visual capacity C_V^T and perceived signal-to-noise ratio S/N as a function of the number of samples for the horizontal direction of a hypothetical television display. The viewing distance is 4 picture heights, and the bandwidth limitation has been set by the U.S. Monochrome Standards (N_{TV} = 471). Curves are shown for s = 0 and s = 1.

the vertical direction there are the customary 480 samples with practically no band limitation on the input signal, so that we have taken $\omega_M = \infty$, corresponding to the extreme undersampled limit. We use the value s = 1, i.e., full-width sampling. In the horizontal direction, instead of the usual analog scanning system, we have taken $N_s = 740$ samples with the band-limiting input characteristic $R(\omega)$ corresponding to either the full U.S. monochrome bandwidth of 4.5 MHz ($N_{TV} = 471$) or a 3.5-MHz ($N_{TV} = 366$) cutoff, in accordance with current U.S. practice. Thus, the picture-producing capability in the horizontal direction is that of an oversampled display, sampled at 1.6 or 2.0 times the Nyquist rate $N_s = N_{TV}$, depending on the particular bandwidth chosen. We employed a constant printing function, corresponding to $\hat{\Pi}(\omega) = \text{sinc}\ (p\omega/\omega_s)$, where p = 1 for complete uniform illumination of one sampling location [Eq. (54)]. We used the value p = 1 in all cases except for the curve labeled "Scanning Beam Width = 2 Scan Lines" for which we took p = 2. It is seen from the figure that the two dashed curves for the 480 samples follow the expected behavior with viewing distance, eventually achieving the r^2 dependence predicted by Eq. (70). The case p = 2 gives superior signal-to-noise performance, particularly at the smaller viewing distances, in agreement with our observation that wide printing functions increase S(r)/N(r). By examining Eqs. (50) to (52) for S(r)/N(r), one can readily show that, for $\hat{\Pi}(\omega) = \text{sinc}\ (p\omega/\omega_s)$ and for p and s both integers, the parameters s and p can be interchanged leaving S(r)/N(r) unchanged. Therefore, the upper dashed curve in Fig. 11 also represents the S(r)/N(r) for p = 1 and s = 2. Recalling that the lower dashed curve is calculated for s = p = 1, we see that this result illustrates the beneficial effects of large sampling widths in undersampled displays. The three solid curves for 740 samples lie well above the curves for 480 samples. Most of the difference between the two sets of curves arises from the fact that the horizontal direction of our hypothetical display is oversampled whereas the vertical direction is undersampled. Very little difference arises from the larger number of samples *per se*, since the 4:3 aspect ratio of the television screen brings the number of samples per unit length for the horizontal and vertical directions to within about 15% of each other. It can be verified that the S(r)/N(r) for the solid curves approaches the characteristic $r^{5/2}$ law, as predicted by Eq. (62). Furthermore, note that raising the sampling width from s = 0 to s = 1 produces very little change of S(r)/N(r), also in

agreement with Eq. (62) and our general observation that s is an important parameter only in undersampled displays. Finally, note that the utilization of the full 4.5-MHz (N_{TV} = 471) U.S. monochrome standard bandwidth gives a somewhat lower S(r)/N(r) than the 3.5-MHz (N_{TV} = 366) system, also in general accord with Eq. (62).

Figure 12 illustrates the effect of varying the width of the printing function on the S(r)/N(r) for the vertical (sampled) direction of television displays. The values of p, the width parameter for the Gaussian printing function Eq. (77), given in the figure are the same as those employed to calculate the corresponding $C_v^T(r)$ shown in Fig. 9. The particular value p = 0.642 was the value measured for the real kinescope described in the discussion of the visual capacity. Thus, the curve for p = 0.642 is actually the same as in Fig. 10. It can be verified from the figure that S/N varies approximately as $\exp(\pi^2 p^2)$ at very small viewing distances, whereas S/N is only weakly dependent on p and approaches the expected r^2 dependence at large viewing distances. All this is in accord with the behavior predicted by the analytic approximations, as summarized in Tables 1 and 2. Comparisor of Figs. 9 and 12 shows a striking example of the rule that narrow-width printing functions favor the visual capacity but only at the expense of decreased perceived signal-to-noise ratio. The optimum tradeoff between visual capacity and perceived signal-to-noise ratio will be discussed in the next section.

As an example of the importance of the sampling rate in determining picture quality, in Figs. 13 and 14 the visual capacity and perceived signal-to-noise ratio, calculated from Eqs. (23), (27), and (50) to (52), are simultaneously plotted against the number of samples at two viewing distances for a display operating with a band-limited input characteristic corresponding to the full U.S. monochrome bandwidth (N_{TV} = 471). The printing function is taken to be that of Eq. (54), i.e., uniform illumination of a single sampling location (p = 1). Curves are shown for both s = 0 and s = 1. We have again taken the television format of a 4:3 aspect ratio, so the viewing distances r = 3h and 4h correspond to 2.25 and 3 picture widths, respectively. At each viewing distance, C_v^T rises with the number of samples, eventually saturating in the analog limit at a value determined by the band-limited characteristics of the display and by the limitations of the visual system. For the U.S. monochrome standards, the maximum value of C_v^T in the analog limit, as calculated from

Eq. (26), is approximately 450 and is achieved at $r_p = 4.03h$. On the other hand, S/N increases without limit, rising rapidly from its low value near the Nyquist sampling rate $N_s = N_{TV} = 471$. Examination of Figs. 13 and 14 shows that very little gain in S/N is achieved by employing full-width sampling $s = 1$ rather than delta-function sampling $s = 0$. This is another example of the analytically derived result that S/N is not sensitive to the value of s in oversampled displays. On the other hand, by using $s = 0$, C_v^T is strongly enhanced in the range of N_s under consideration. This result follows from Eq. (27) for C_v^T. Therefore, our results strongly indicate that narrow-width sampling is preferred on an overall basis in oversampled displays.*

From the $s = 0$ curves in Figs. 13 and 14, one sees that, at a viewing distance of 3h, nearly 700 samples are required to give S/N = 30, a satisfactory value. At this point, C_v^T is about 340, and a gain of less than 10% in C_v^T is obtained by a 50% increase of the number of samples to 1050. At a viewing distance of 4h, only about 570 samples are required to achieve S/N = 30. For 570 samples and $s = 0$, $C_v^T \approx 360$, and a 17% increase is achieved by going to 1050 samples. Thus, almost identical performance in terms of sharpness and perceived signal-to-noise ratio is obtained with 570 samples at a viewing distance of 4h as is found with nearly 700 samples at a viewing distance of 3h. This simple example serves to illustrate the kind of tradeoffs that one can consider when the human observer is explicitly included in the calculation of performance criteria.

e. The Correlation Quality and the Mean Square Perceived Error. - In closing this section, we derive an important relation between the perceived signal and noise powers and two new quantities which have interesting and useful properties. The first of these new quantities is the *correlation quality* [17] Q defined by

$$Q = \frac{\overline{\langle E(x)E_o(x)\rangle}}{\overline{\langle E_o^2(x)\rangle}} \tag{79}$$

*This result does not take into consideration possible beneficial effects of large values of s in integrating out electronic noise that may be present along with the video signal before the sampling process.

where $E_o(x)$ is the perceived intensity pattern reproduced by a perfect display, i.e.,

$$E_o(x) = \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi}\, 0(\omega r/2\pi)\, \hat{I}_o(\omega)\, \exp(i\omega x) \tag{80}$$

The quantity Q is proportional to the cross correlation function between the perceived scene as actually reproduced by the display and the scene produced by a perfect reproduction device. It is, therefore, a measure of the degree to which the perceived picture resembles a perfect reproduction of the original scene. Making use of the theorem of Section III.C [Eq. (22)], Eq. (79) becomes

$$Q = \frac{\overline{\langle E_S(x)\, E_o(x)\rangle}}{\overline{\langle E_o^2(x)\rangle}} \tag{81}$$

The second quantity is the *mean square perceived error* [17] ε, defined by

$$\varepsilon = \frac{\overline{\langle (E(x) - E_o(x))^2\rangle}}{\overline{\langle E_o^2(x)\rangle}} \tag{82}$$

From this definition, it is evident that ε is the mean square deviation of the perceived picture from the perfect reproduction, normalized to $S_o^2 = \langle E_o^2(x)\rangle$, the perceived signal power for the perfect display. To obtain the desired relationship, we expand Eq. (82) and obtain

$$\varepsilon = 1 - 2Q + \overline{\langle E^2(x)\rangle}/\overline{\langle E_o^2(x)\rangle} \tag{83}$$

where we have used Eq. (79) for Q. Next, we note that S_o^2 can be obtained from Eq. (50) by setting $R_{eff}(\omega) = 1$. Then, using Eq. (47) to express $\overline{\langle E^2(x)\rangle}$ in terms of the signal power S^2 and the noise power N^2, Eq. (83) becomes

$$\varepsilon = 1 + 2Q + (S^2 + N^2)/S_o^2 \tag{84}$$

where

$$S_o^2(r) = \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi} O^2(\omega r/2\pi)\ \Phi(\omega) \tag{85}$$

Eq. (84) is a relation between the correlation quality, the mean square perceived error, the average perceived signal powers for the actual and perfect displays, and the average perceived noise power.

We can write Eq. (84) in explicit form using the following expression for Q(r):

$$Q(r) = \frac{\int_{-\infty}^{+\infty} \frac{d\omega}{2\pi} O^2(\omega r/2\pi)\ \mathrm{Re}\ R_{eff}(\omega)\, \Phi(\omega)}{\int_{-\infty}^{+\infty} \frac{d\omega}{2\pi}\ O^2(\omega r/2\pi)\ \Phi(\omega)} \tag{86}$$

Equation (86) is easily proved using the technique of Section III.C. Taking the expectation value of $E_S(x)\ E_o(x)$ naturally introduces the factor $R_{eff}(\omega)$, the effective MTF that generates the signal part $E_S(x)$ of the total perceived response E(x). Then, because of the reality of $E_S(x)$, the integration over all ω projects out only the real part of $R_{eff}(\omega)$. Substituting Eqs. (85) and (86) into Eq. (84) and making use of Eqs. (50) and (51) for S^2 and N^2, respectively, gives the expression for $\varepsilon(r)$,

$$\varepsilon(r) = \frac{\int_{-\infty}^{+\infty} \frac{d\omega}{2\pi} O^2(\omega r/2\pi) \left\{ [1-R_{eff}(\omega)][1-R^*_{eff}(\omega)]\ \Phi(\omega) + |R_{eff}(\omega)|^2 N(\omega) \right\}}{\int_{-\infty}^{+\infty} \frac{d\omega}{2\pi} O^2(\omega r/2\pi)\ \Phi(\omega)} \tag{87}$$

From Eqs. (86) and (87) with $\Phi(\omega) = B/\omega^2$, it is a simple matter to show: (1) $\lim_{r\to\infty} Q(r) = 1$ (an imperfect picture cannot be distinguished from a perfect one at very large viewing distances), (2) $\lim_{r\to 0} Q(r) = 0$ (there is no correlation between the perceived picture and a perfect picture at very small viewing distances, a consequence of the band limitation of all displays), (3) $\lim_{r\to\infty} \varepsilon(r) = 0$

(there is no perceived error at large viewing distances, a result consistent with (1) above), and (4) $\lim_{r \to 0} \varepsilon(r) = 1 + N^2(r)/S_o^2(r)$ (the normalized mean square error is equal to unity plus a term proportional to the noise power). Figure 15 illustrates this behavior for the case of noiseless analog television displays operating with a 3.5-MHz band-limited input characteristic. Curves of $\varepsilon^{1/2}$ and Q are shown as a function of viewing distance for an ideal kinescope $R_o(\omega) = 1$ and for the actual measured MTF of a color kinescope*. The vertical arrows in the figure indicate the calculated viewing distance r_p for maximum visual capacity. It is interesting to note that, at the viewing distance for maximum edge appreciation, the correlation quality is greater than 0.9 and the rms perceived error is in the range 0.16 to 0.18.

E. THE TOTAL INFORMATION CAPACITY - A COMBINED PERCEIVABLE INFORMATION DESCRIPTOR

In Section III.D, we described how one can characterize a sampled display in terms of visual capacity (its "sharpness" or edge discrimination ability in the absence of noise) and the perceived signal-to-noise ratio (a measure of the amount of disturbing noise, relative to the perceived signal content). We presented evidence that values of the perceived signal-to-noise ratio in the neighborhood of 20 are representative of high-quality television displays, viewed at normal distances. We indicated how one can quantify the effect of varying such parameters of the display process as bandwidth, number of samples, sampling width, and printing width. It was indicated that questions regarding the optimization of the parameters of the sampling process involve the tradeoff between edge discrimination ability and perceived signal-to-noise ratio. An example where such a tradeoff must be considered is the case of the effective printing width p; small values of p favor visual capacity but at the expense of the perceived signal-to-noise ratio, and vice versa. Therefore, a single descriptor that combines both visual capacity and perceived signal-to-noise ratio would be of obvious utility. For example, suppose we are faced with the question of what is the optimum printing width at a particular viewing distance, given the constraints that total luminance (brightness times area) and all other properties of the sampling process are held constant.

*E. W. Herold, private communication.

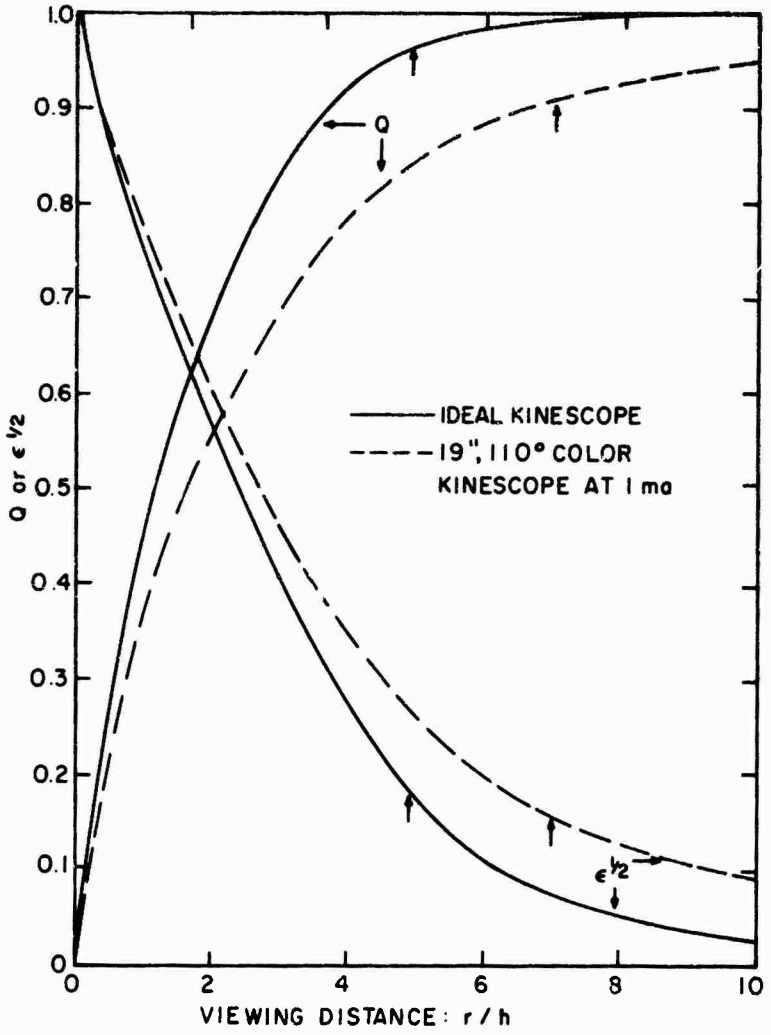


Figure 15. Correlation quality Q and the perceived root mean square error $\varepsilon^{1/2}$ as a function of viewing distance in picture heights for an ideal noiseless kinescope and for a commercially available kinescope (assumed to be noiseless) at the indicated value of the beam current. A band-limited input characteristic with a cutoff of 3.5 MHz (N_{TV} = 366) was employed in the calculations. The vertical arrows indicate the calculated viewing distances r_p for maximum visual capacity.

We might guess that at very small viewing distances, it would be beneficial to employ large values of p in order to avoid as much as possible the large amounts of disturbing noise introduced by narrow printing functions (see, for example, Fig. 12). On the other hand, at very large viewing distances, the eye would be unable to perceive this noise, and a continuous picture would be observed. In that case, it might be beneficial to use the smallest practical active area (again keeping the total number of photons constant) in order to gain a sensation of sharpness from the finer printing (see, for example, Fig. 9). If this line of reasoning is correct, there exists an optimum value of p which depends explicitly on viewing distance and the other parameters of the sampling process. A unified descriptor should be capable of predicting this optimum value.

A unified descriptor must weigh C_v^T and S/N in a manner that parallels, as closely as possible, the way in which the human observer weighs the relative virtues of sharpness and freedom from disturbing noise. We have no *a priori* knowledge and little *a posteriori* knowledge in this area, so that any approach must be, to a certain degree, *ad hoc*. However, we can heuristically continue our approach based on statistical communication theory and consider the total information capacity of the display-observer system.

It is well-known [10, 18, 19] that the limiting rate of information transfer of a communication system is determined by the system's bandwidth and its overall signal-to-noise ratio. The bandwidth determines the maximum rate of transmission μ of the information-carrying pulses, and the signal-to-noise ratio fixes the number of levels q which each pulse can assume. For a communication channel that can handle q pulse levels and whose maximum transmission rate is μ pulses/second, the system can assume $S = q^{\mu T}$ discrete states during a message duration time T. The information capacity of the channel is $H = \log_2 S = \mu T \log_2 q$ and is measured in bits.

In real communication systems, the maximum number of permissible levels is determined by the available signal-to-noise ratio, the acceptable transmission error rate, and the details of the encoding and decoding methods employed. However, Shannon [10] has shown that the maximum information that can be transmitted in time T without error is given by

$$H = \Omega T \log_2 \left[1 + (S/N)_T^2\right] \tag{88}$$

where $(S/N)_T^2$ is the total system signal-to-noise ratio for power, and Ω is the bandwidth of the system. We may regard Eq. (88) as arising from the replacement of the pulse transmission rate [18, 19] by twice the bandwidth and the number of pulse levels q by the quantity [18, 19] $[1 + (S/N)_T^2]^{1/2}$.

A measure of the total information capacity of the communication channel consisting of display and observer can be obtained by defining H as the maximum amount of visual information that a human observer can perceive across a display of width w. Recalling that the visual capacity C_v^T is defined as the number of fully resolvable edge transitions that can be perceived across a display, it is natural to identify the bandwidth Ω in Eq. (88) as one-half the visual capacity. Thus, we have

$$\Omega = \frac{1}{2} C_v^T \tag{89}$$

Next, we consider the total system signal-to-noise ratio $(S/N)_T$. In our display-observer communication system, noise arises from both the display and the observer. The display noise power N^2 represents noise as it is produced by the picture reproduction device and filtered by the MTF of the human visual system. The observer noise N_v^2 is a characteristic of the human visual system. As these noise powers are uncorrelated, it is reasonable to assume that they are additive

$$N_T^2 = N^2 + N_v^2 \tag{90}$$

Next, we make the simplifying assumption, consistent with experimental observation, that the noise N_v^2 of the human visual system is set by the signal level; the noise rides with the signal so as to maintain a constant effective signal-to-noise ratio of the human visual system $(S/N)_h$, regardless of the signal level.* This means that we may write

$$N_v^2 = S^2 (S/N)_h^{-2} \tag{91}$$

---

*This assumption is consistent with the Weber-Fechner law, which states that the minimum observable brightness difference is proportional to the brightness [14].

whereupon Eqs. (90) and (91) give the following formula for the total system signal-to-noise ratio

$$(S/N)_T^2 = \left[(S/N)^{-2} + (S/N)_h^{-2}\right]^{-1} \tag{92}$$

Here $(S/N)^2$ is the perceived signal-to-noise ratio, given by Eqs. (50) and (51). As a function of S/N, the total system signal-to-noise ratio is equal to S/N for $(S/N)^2 << (S/N)_h^2$ but eventually saturates at the value $(S/N)_h$ as S/N is increased indefinitely. This behavior is in qualitative agreement with our experience in examining noisy pictures. We know that beyond a certain point, there is no advantage to be gained in increasing the display signal-to-noise ratio. The improvements are simply not perceived. On the basis of (1) our observation in Section III.D that a high-quality television display is characterized by values of S/N in the neighborhood of 20, and (2) the Weber-Fechner law, we expect $(S/N)_h$ to be on the order of 10. This value of $(S/N)_h$ implies that the maximum number of perceivable levels q for the human visual system in the absence of display noise is approximately 10. However, we know that far more grey-scale levels are perceivable in a typical display with, say, a maximum contrast ratio of 100:1. The reason for this discrepancy lies in our treatment of the human visual system as a linear system, whereas the Weber-Fechner law establishes the position of perceivable gray-scale levels exponentially. The assignment of a single number to the effective signal-to-noise ratio of the human visual system is undoubtedly an over-simplification, since it may change with viewing distance and other display and environmental variables, but in our simple heuristic approach, we shall take $(S/N)_h$ to be a single parameter. Numerical calculations have indicated that, within the expected range of values of $(S/N)_h$, conclusions derived from the calculations are insensitive to the precise value of this parameter.

Having determined the quantities corresponding to the bandwidth and the total system signal-to-noise ratio for the case of a display-observer system, we can substitute Eqs. (89) and (92) for these quantities into Eq. (88) for the total information capacity H. Making the dependence on viewing distance explicit, we have

$$H(r) = \frac{1}{2} C_v^T(r) \log_2 \left\{1 + [(S(r)/N(r))^{-2} + (S/N)_h^{-2}]^{-1}\right\} \tag{93}$$

Even though the basic assumptions used in deriving Eqs. (92) and (93) are consistent with experimental observations, they cannot be directly verified through state-of-the-art psychophysical measurements. Nevertheless, Eqs. (92) and (93) can be arrived at through purely formalistic arguments. The physical significance of the signal-to-noise ratio is that it determines the number of distinguishable levels q. We know from experimental observations (the Weber-Fechner law) that there exists in the human visual perception process some mechanism that sets the minimum discernable contrast difference even in the absence of image noise. (We assume that the image is sufficiently bright that quantal fluctuations [14] are not dominant.) We can describe this internal mechanism by an effective signal-to-noise ratio $(S/N)_h$ such that for noise-free images, the number of discernable levels is $q = [1 + (S/N)_h^2]^{1/2}$. If the perceived image signal-to-noise ratio $S(r)/N(r)$ is much poorer than $(S/N)_h$, then the number of levels is determined by the image noise, so that $q = [1 + (S(r)/N(r))^2]^{1/2}$. The above two expressions for q describe the observer-display system's behavior in the asymptotic regimes of $S(r)/N(r) \to \infty$ and $S(r)/N(r) \to 0$, respectively. In the range where $S(r)/N(r) \cong (S/N)_h$, we expect that the number of discernable levels is smaller than that predicted by either one of the above asymptotic formulas. In the absence of additional information, we now wish to combine mathematically the effects of the internal mechanism and of the image noise in such a way that the above expectation and the asymptotic behaviors are properly described. The simplest mathematical expression that satisfies these requirements is obtained by defining a total effective signal-to-noise ratio $(S/N)_T$ through Eq. (92). Substituting Eqs. (89) and (92) into Eq. (88), we arrive at the expression for H(r) given in Eq. (93).

Equation (93), along with Eq. (26) or (27) for $C_v^T(r)$ and Eqs. (50) and (51) for $S^2(r)$ and $N^2(r)$, constitutes our proposal for a unified descriptor. It should apply to sampling noise, for which the noise power spectrum is given in Eq. (52), and other forms of noise for which the noise power spectrum is specified. The quantity H represents the total information capacity of the display-observer system, including the effects of both edge discrimination ability and noise perception. By analogy with ordinary communication channels, H is to be measured in bits. It should be emphasized that H in no way represents the actual information transfer from the original scene to the perceptual level.

a. Properties of the Total Information Capacity. - The properties of H(r) are easily derived from those of $C_v^T(r)$ and S(r)/N(r). In the noiseless display limit $[S(r)/N(r)]^2 >> (S/N)_h^2$, the perceived signal-to-noise ratio does not affect the total information capacity, so the H(r) is equal to a constant times $C_v^T(r)$:

$$H(r) = \frac{1}{2} C_v^T(r) \log_2 \left\{ 1 + (S/N)_h^2 \right\}; \text{ for } [S(r)/N(r)]^2 >> (S/N)_h^2 \quad (94)$$

As S(r)/N(r) is decreased, the perceived noise decreases H(r) from the noiseless display limit Eq. (94). Indeed in the limit $[S(r)/N(r)]^2 << (S/N)_h^2$, the total system signal-to-noise ratio is determined by the display noise, and H(r) is reduced in a logarithmic manner from the noiseless display limit:

$$H(r) = \frac{1}{2} C_v^T(r) \log_2 \left\{ 1 + (S(r)/N(r))^2 \right\}; \quad \text{for } [S(r)/N(r)]^2 << (S/N)_h^2 \quad (95)$$

In Section III.D, we showed that, for sampled displays, S(r)/N(r) is a monotonically increasing function of r, rising from a finite or zero value at r = 0 and diverging as $r \to \infty$ (see Tables 1 and 2). On the other hand, $C_v^T(r)$ exhibits a peak at a viewing distance r_p and approaches zero at both r = 0 and $r \to \infty$. From these results, it is clear that H(r) will also always exhibit a peak at a particular viewing distance which is $\geq r_p$. We call the viewing distance for maximum H(r) the optimum viewing distance and denote it by r_{opt}. We do so on the grounds that r_{opt} represents the viewing distance which maximizes the total information capacity of the display-observer system, including the effects of both edge discrimination ability and perceived noise.

b. Sharpness-Limited and Noise-Limited Displays. - We find it convenient and natural to divide displays into two categories:

(1) Sharpness-limited displays: those displays for which S(r)/N(r) has risen to a value greater than $(S/N)_h$ at the viewing distance r_p for maximum $C_v^T(r)$. An observer located at $r = r_p$ does not find noise objectionable.

(2) Noise-limited displays: those displays for which S(r)/N(r) is considerably less than $(S/N)_h$ at r_p. Noise is very apparent to an observer located at r_p.

For sharpness-limited displays, perceived noise does not appreciably affect the value of H(r) for $r \gtrsim r_p$. Thus, Eq. (94) holds in this viewing distance range, and $r_{opt} \cong r_p$. An example of a sharpness-limited display is the vertical (sampled) direction of the television display whose $C_v^T(r)$ and S(r)/N(r) are graphed in Figs. 8 and 10, respectively. The calculated $S(r_p)/N(r_p) = 21$ at the viewing distance $r_p = 4.7h$ puts this display in the sharpness-limited class. Using the calculated $C_v^T(r)$ and S(r)/N(r), we obtain from Eq. (93), with $(S/N)_h = 10$, the graph of H(r) indicated by the dashed curve in Fig. 16. For comparison, we also show the calculated H(r) for the horizontal analog scanning direction of this display, using the $C_v^T(r)$ from Fig. 8 and assuming noiseless picture reproduction [S(r)/N(r) = ∞]. Comparing Fig. 16 with Fig. 8, it is seen that the effect of the sampling noise on the total information capacity is minor for $r \gtrsim 3h$. The optimum viewing distance is $r_{opt} = 5.1h$ instead of the value $r_p = 4.7h$ for maximum visual capacity. The decrease of the peak value of H(r) due to sampling noise is only about 4% of the value $\frac{1}{2} C_v^T(r_p)\log_2[1 + (S/N)_h^2] = 929$ bits that would have been achieved in the absence of sampling noise. For $r < 3h$, the H(r) for the vertical direction drops much more abruptly than that for the horizontal direction, actually falling below the latter curve for $r \leq 1.3h$. This effect is due to the rapidly decreasing S(r)/N(r), as shown in Fig. 10, which drastically reduces H(r) at small viewing distances.

In the case of noise-limited displays, the effect of the noise is to reduce H(r) appreciably in the neighborhood of r_p. It then becomes advantageous to increase the viewing distance in order to raise the perceived signal-to-noise ratio. The distance r_{opt} is then determined by a tradeoff between the decrease of $C_v^T(r)$ at large viewing distances, as given by Eq. (28), and the natural increase of S(r)/N(r). As an illustrative and significant example, we consider the case of an analog display with a flat overall MTF $R_o(\omega) = 1$ for $|\omega| < \omega_M$. We assume a white noise power spectrum $N(\omega) = N(0)$ so that the noise power N_s^2, as measured on the display screen, is given by Eq. (53):

$$N_s^2 = (\omega_M/\pi)\, N(0) \tag{96}$$

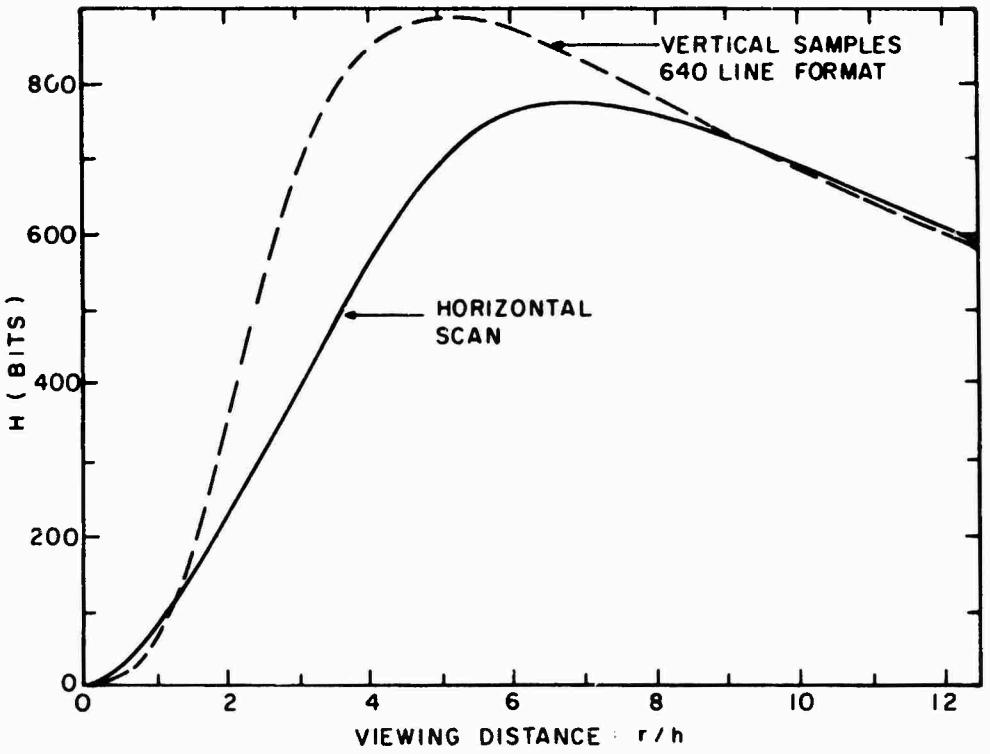


Figure 16. Total information capacity H as a function of viewing distance in picture heights for a real television display device. The H for both vertical and horizontal directions are shown. The value of the parameter $(S/N)_h$ was taken to be 10. An N_s = 640-line format has been employed in order to compensate for the 4:3 aspect ratio of the display screen. The C_v^T and S/N for this display are given in Figs. 8 and 10, respectively.

The perceived noise power $N^2(r)$ for constant $N(\omega)$ is obtained from Eq. (51). Setting $R_{eff}(\omega) = 1$ for $|\omega| \leq \omega_M$, we have

$$N^2(r) = (\pi N_s^2/\omega_M) \int_{-\omega_M}^{+\omega_M} \frac{d\omega}{2\pi} O^2(\omega r/2\pi) \tag{97}$$

We are interested in the viewing distance regime $r >> r_p$, where, according to Eq. (34), r_p is approximately $2\pi\nu_o/\omega_M$. In that case, the limits on the integration in Eq. (97) may be extended to $\pm\infty$, giving us the following expression for $N^2(r)$:

$$N^2(r) = (\pi/\omega_M \Theta_e^\infty r)\ N_s^2;\ \text{for } r >> 2\pi\nu_o/\omega_M \tag{98}$$

Here we have employed Eq. (29) to express the integral in terms of θ_e^∞, the perceived angular width of a single edge transition. Turning to the perceived signal power, for the measured signal power spectrum $\Phi(\omega) = B/\omega^2$, we can make use of Eqs. (55) and (63) for $S^2(r)$ in the large viewing distance limit. We have then

$$S^2(r) = (4.73\ Br/2\pi^2\nu_o);\ \text{for } r >> 2\pi\nu_o/\omega_M \tag{99}$$

Next, we go back to Eq. (42b) in order to express the power amplitude coefficient B in terms of the ensemble average input power $\langle I_o^2(x)\rangle$. For $\omega_M >> \omega_L$, the average signal power S_s^2, as measured on the display screen, is identical to the average scene power within the field-of-view of the imaging system, $\langle I_o^2(x)\rangle$. This is because the $1/\omega^2$ power spectrum concentrates most of the signal power at low frequencies, thereby rendering S_s^2 independent of ω_M. With these considerations, Eq. (99) becomes

$$S^2(r) = (4.73\ \omega_L r/\pi^2\nu_o)\ S_s^2;\ \text{for } r >> 2\pi\nu_o/\omega_M \tag{100}$$

Combining Eqs. (98) and (100), we obtain the asymptotic behavior of $S(r)/N(r)$ for white noise at large viewing distances:

$$S(r)/N(r) = 1.88 \times 10^{-3}\ N_{TV}^{1/2}(r/w)\ (S_s/N_s);$$
$$\text{for } r/w >> 900/N_{TV} \tag{101}$$

In arriving at Eq. (101), we have used the value of θ_e^∞ given in Eq. (30), employed the expression $N_{TV} = \omega_M\omega/\pi$ to write the maximum frequency in terms of the maximum number of TV lines, and once again set $\omega_L = 2\pi/w$ and $\nu_o = 458$ cycles/radian-of-vision. From Eq. (101), it is seen that $S(r)/N(r)$ is proportional to S_s/N_s, the signal-to-noise ratio as measured on the display screen, and also varies linearly with viewing distance. The factor $N_{TV}^{1/2}$ arises because the assumed white noise spectrum gives an rms noise fluctuation N_s that is proportional to the square root of the bandwidth [see Eq. (96)]. It is convenient to think of the quantity $N_s/N_{TV}^{1/2}$ in Eq. (101) as representing $[N(0)]^{1/2}$, the square root of the noise spectral density.

Given Eq. (101) for $S(r)/N(r)$ and Eq. (28) for $C_v^T(r)$ in the large viewing distance limit, we differentiate Eq. (93) for H(r) in order to calculate the optimum viewing distance. The operation is straightforward and gives us the following condition for r_{opt}:

$$\left[(S/N)^2 \frac{\left\{1 + (S/N)_T^2\right\} \ln \left[1 + (S/N)_T^2\right]}{(S/N)_T^4} \right] \Bigg|_{r = r_{opt}} = 2 \tag{102}$$

where $(S/N)_T$ is the total system signal-to-noise ratio, given in Eq. (92). Equations (92) and (102) can be solved numerically for the required value of $S(r_{opt})/N(r_{opt})$. We find that $S(r_{opt})/N(r_{opt})$ is essentially independent of the value of $(S/N)_h$, varying from the value 1.90 for $(S/N)_h = 10$ to 1.98 for $(S/N)_h = \infty$. We take the former value, thereby giving us the result that, for extremely noisy displays, the viewing distance for maximum H must be increased from r_p until the condition

$$S(r_{opt})/N(r_{opt}) = 1.90 \tag{103}$$

is met. From Eq. (101), it is seen that Eq. (103) is satisfied when

$$r_{opt}/w = 1010/[N_{TV}^{1/2} (S_s/N_s)] \tag{104}$$

Since we demand that $r_{opt} >> r_p$, Eq. (104) can be used to establish a condition on the value S_s/N_s. From the inequality in Eq. (101), we have

$$S_s/N_s << N_{TV}^{1/2} \tag{105}$$

The maximum value of H(r), corresponding to the viewing distance given in Eq. (104), is easily obtained from Eq. (93) with $C_v^T(r) = w/\theta_e^\infty r$. Taking $(S/N)_h = 10$, we compute

$$H(r_{opt}) = 2.0\ N_{TV}^{1/2} (S_s/N_s) \tag{106}$$

It is instructive to cast this result in terms of the maximum visual capacity of a noiseless analog display with the same value of $H(r_{opt})$. In order to accomplish this, we use the fact that a display with a flat overall MTF $R_o(\omega) = 1$ for $|\omega| \leq \pi N_{TV}^{eq}/w$ has a maximum visual capacity [11] $C_v^T(r_p) = N_{TV}^{eq}$. Then from Eqs. (94), (105), and (106) with $(S/N)_h = 10$, we see that an extremely noisy display with a white noise spectrum has the same maximum total information capacity as a noiseless display with a limiting resolution given by

$$N_{TV}^{eq} = 0.60\, N_{TV}^{1/2}\, (S_s/N_s);\ \text{for}\ N_{TV}^{eq} << N_{TV} \tag{107}$$

Equation (107) is remarkable, for it says that an extremely noisy display is equivalent in total information capacity to a noiseless analog display whose resolution is proportional to the signal-to-noise ratio, as measured on the screen of the noisy display. This is true despite the fact that H(r) depends logarithmically on the total system signal-to-noise ratio. Furthermore, using Eqs. (42) and (96), it is possible to eliminate the factor $N_{TV}^{1/2}$ on the right-hand side of Eq. (107) and express N_{TV}^{eq} entirely in terms of the ratio of the signal power amplitude coefficient B to the noise power spectrum N(0). Thus, the total information capacity of an extreme noise-limited display is independent of its resolution. The display is so noisy that its optimum performance depends only on the input signal power and the noise power spectrum. This situation is analogous to the case of an ordinary noisy communication channel. Suppose, for example, that we are receiving information from a distant source such as a space probe. As the source recedes, the intrinsic signal power decreases, while the intrinsic noise power remains constant. If we want to keep the error rate constant, we must increase the signal integration time proportionately in order to maintain the required overall signal-to-noise ratio. However, this means that the bandwidth, and hence the information transmission rate, is decreased by the same factor as the signal power. In effect, the information transmission rate is proportional to the ratio of the signal power to the noise power. For the case of the display-observer communication channel, increasing the viewing distance is analogous to increasing the integration time. As the ratio of signal power to noise power decreases, the observer must increase his viewing distance in order to maintain the value

of the perceived signal-to-noise ratio given by Eq. (103). As the viewing distance increases, the "bandwidth" or visual capacity decreases proportionately [Eq. (28)]. The net effect is that the maximum information capacity is proportional to the ratio of signal power to noise power.

c. The Equivalence Factor and its Relation to the Kell Factor. - The concept of an equivalence factor that can be employed to convert the overall picture reproduction ability of a sampled display with N_s samples into an effective number of TV lines for an analog display has had historical appeal. The early work of Kell [20] and coworkers showed empirically that, for television displays, the picture reproduction ability of N_s samples is equivalent to an analog scanning display system with a limiting resolution corresponding to $N_{TV} = K N_s$, where K is the so-called Kell factor. Values of K in the range 0.53 to 0.85 were observed [21].

The formalism presented in this report allows one to calculate the required number of TV lines N_{TV} for a noiseless analog display to produce the same maximum total information capacity as a sampled display with N_s samples. We define the equivalence factor & according to

$$\& = N_{TV}/N_s;\ H(r_{opt})\ |_{N_{TV}} = H(r_{opt})|_{N_s} \tag{108}$$

where it is understood that N_{TV} and N_s give the same value of H(r) at their respective optimum viewing distances. It is clear that & will depend on the various parameters of the sampled display (sampling width, printing width, bandwidth) and the characteristics assumed for the analog display [the form of the overall MTF $R_o(\omega)$]. As an example, we consider the vertical (sampled) direction of the television display whose H(r) is indicated by the dashed curve in Fig. 16. From the figure, we see that $H(r_{opt}) = 890$ bits. According to Eq. (94) for H(r) in the absence of display noise, this value of $H(r_{opt})$ is also obtained for a noiseless analog display with a maximum visual capacity $C_v^T(r_p) = 2 \times 890/\log_2[1 + (S/N)_h^2] = 267$ transitions. This value of $C_v^T(r_p)$ can be achieved by an infinite number of analog displays, depending on the specific form of the overall MTF $R_o(\omega)$. If & is to have any practical meaning, we should confine ourselves to the $R_o(\omega)$ for the horizontal scanning direction of the same television display device. We see from Fig. 8 that, with the contribution of the video circuit response function to $R_o(\omega)$ used in the

calculation of $C_v^T(r)$ for the horizontal direction, the maximum C_v^T is only about 235 transitions. In order to increase this value to 267 transitions, we find that it would be necessary to extend the cutoff frequency for horizontal signals from $N_{TV} = 366$ to about $N_{TV} = 400$ lines. Thus, in this case, the equivalence factor & is approximately 400/640 = 0.63.

d. Approximate Determination of $(S/N)_h$. - We have argued that the value of the parameter $(S/N)_h$ is expected to be on the order of 10. An attractive means of experimentally determining the value of $(S/N)_h$ that is operationally effective within the context of the present formalism is to assume that the calculated value of r_{opt} for sampled displays or noisy analog displays corresponds to the preferred viewing distance of human observers. Then, the experimental value of $(S/N)_h$ is obtained by varying the value of this parameter until the calculated r_{opt} agrees with the most frequently occurring experimental value. Self-consistency is to be achieved if the same value of $(S/N)_h$ is obtained, within experimental error, from experiments with different sampling parameters and with different scene contents.

For a preliminary experiment, we selected the sampled and unfiltered picture of Lincoln reproduced [4] on the cover of *Science*, June 15, 1973. This picture is characterized by $N_s = 14$ horizontal samples across 2.25 in, extreme undersampling $\omega_M = \infty$, full-width sampling s = 1, and a constant printing function illuminating one entire sampling location [Eq. (54)]. The picture and its immediate surround were illuminated with tungsten lamps to an average brightness of 150 ft-L. The viewing area was approximately 100 ft long and free of obstructions or other known possible bias influences. It was illuminated with natural light to a brightness of over 300 ft-L. The high brightness levels were chosen so as to nearly reproduce the conditions under which the MTF $O(\nu)$, employed in the calculations, was measured [6]. Thirty-six subjects were chosen from the Laboratories population. These included technical and nontechnical, male and female personnel. All subjects with visual defects used corrective lenses. Each subject was asked to choose the distance at which picture looked best to him. The subjects were encouraged to range over as large an excursion of viewing distances as possible before making a decision. No description or commentary was given. The result of the experiment is summarized by the histogram in Fig. 17, which shows that, with a 3-ft

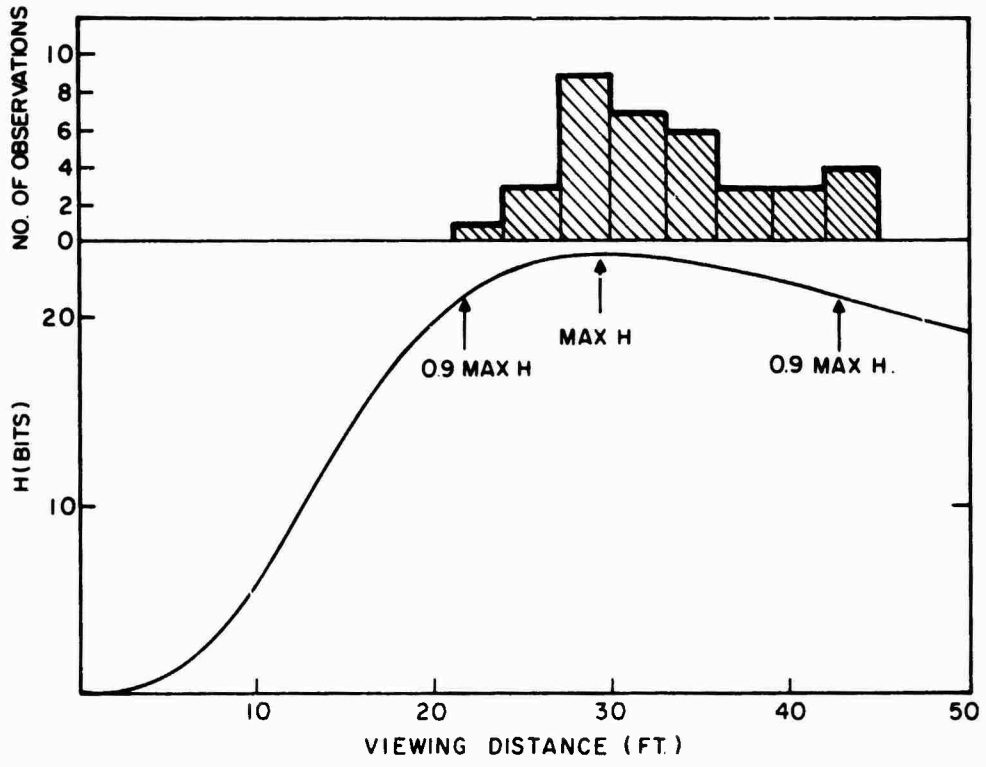


Figure 17. Summary of experimental results and analysis of the sampled "Lincoln" picture [4]. The histogram indicates the number of observations recorded as a function of preferred viewing distance (see text for experimental details). The solid curve represents the total information capacity H for the horizontal direction (N_s = 14) as a function of viewing distance, using the value $(S/N)_h$ = 10. The vertical arrows represent the viewing distance for maximum H and for 90% maximum H.

distance bin, the most frequently occurring value of the preferred viewing distance fell in the range 27 to 30 ft. As can be seen from the figure, a considerable spread was observed; the lowest preferred viewing distance was 22 ft and the largest was 44 ft.

The calculated H(r) for this display, using Eq. (93) with $(S/N)_h$ = 10, is also shown in Fig. 17. It is seen that the peak of H(r) at r_{opt} = 29.3 ft corresponds to the most frequently observed preferred viewing distance. Furthermore, the general shape of the histogram, i.e., skewed to the right, is consistent with the curve of H(r), which rises rapidly at small viewing distances and fal off more gradually at large viewing distances. Note that nearly all observers had a preferred viewing distance that corresponded to more than 90% of the calculated maximum value of H(r).

Although the value $(S/N)_h = 10$ gives an accurate representation of the subjects' preferences, statistical uncertainties in the data and uncertainties in the measured [6] $O(\nu)$ permit a fairly large range of acceptable values of $(S/N)_h$. Taking $(S/N)_h = 15$ would shift r_{opt} to about 32 ft; this value is nearly as valid as that for $(S/N)_h = 10$ because of statistical uncertainties arising out of the relatively small population of subjects. A far more sensitive determination of $(S/N)_h$ could be made from more noise-limited displays, such as one employing delta-function sampling ($s = 0$) and delta-function printing ($p = 0$). Such experiments should be performed.

In spite of these uncertainties, we feel that the results of the experiment and the calculation offer preliminary evidence that the subjects chose the preferred viewing distance on the basis of a compromise between edge discrimination ability and perceived signal-to-noise ratio, in accordance with Eq. (93). The value $(S/N)_h = 10$, derived from the experiment, lies well within the range of values expected on the basis of the Weber-Fechner law and on the basis of the calculated S/N of other noisy displays (see Section III.D). It is not likely that visual capacity alone was operative, since in this case $C_v^T(r)$ has a calculated maximum value at $r_p = 23$ ft. Thirty-five of thirty-six subjects preferred a larger viewing distance. Nor can perceived signal-to-noise ratio alone have been responsible for the observed preferred viewing distances, since such a strategy would favor an infinite viewing distance.

e. The 17-Element Display. - Figures 18 through 22 give the results of calculations appropriate for a 17-element optical block processor of the type described in Section V. In particular, for an extremely undersampled 17-element display, the signal-to-noise ratio S/N(r), the visual capacity $C_v^T(r)$, and H were calculated* for the two extreme cases of delta-function sampling and delta-function printing, and full-width sampling and full-width printing. The properties of these two cases are distinctive and illustrate well our method of analysis.

Figure 18 shows the perceived frequency spectra for the signal and the sampling noise components of the total perceived intensity distribution at a

*For $N_s = 17$, the requirement $\omega_s >> \omega_L = 2\pi/w$, assumed in all the derivations is not extremely well satisfied. Nevertheless, Eqs. (27) and (50) to (52) can be employed to calculate approximately the desired performance characteristics.

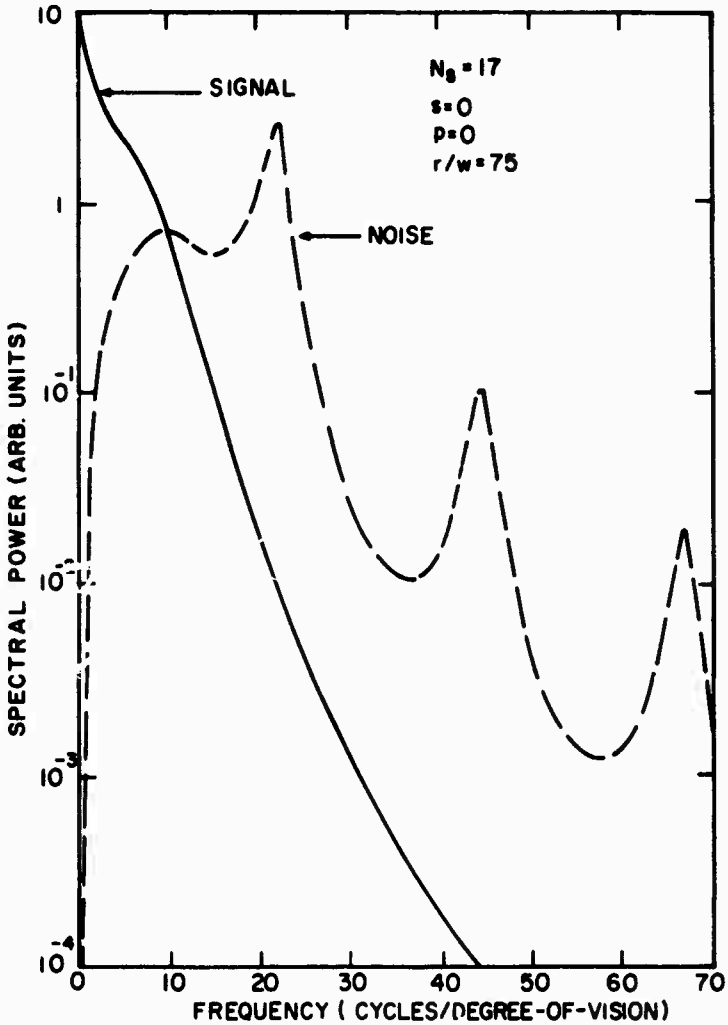


Figure 18. Perceived signal and noise spectra as a function of retinal frequency for a display with N_s = 17 samples at a viewing distance of 75 picture widths. The ordinate represents the integrand of Eqs. (50) and (51) for S^2 and N^2, respectively. The values of s and p are given in the figure.

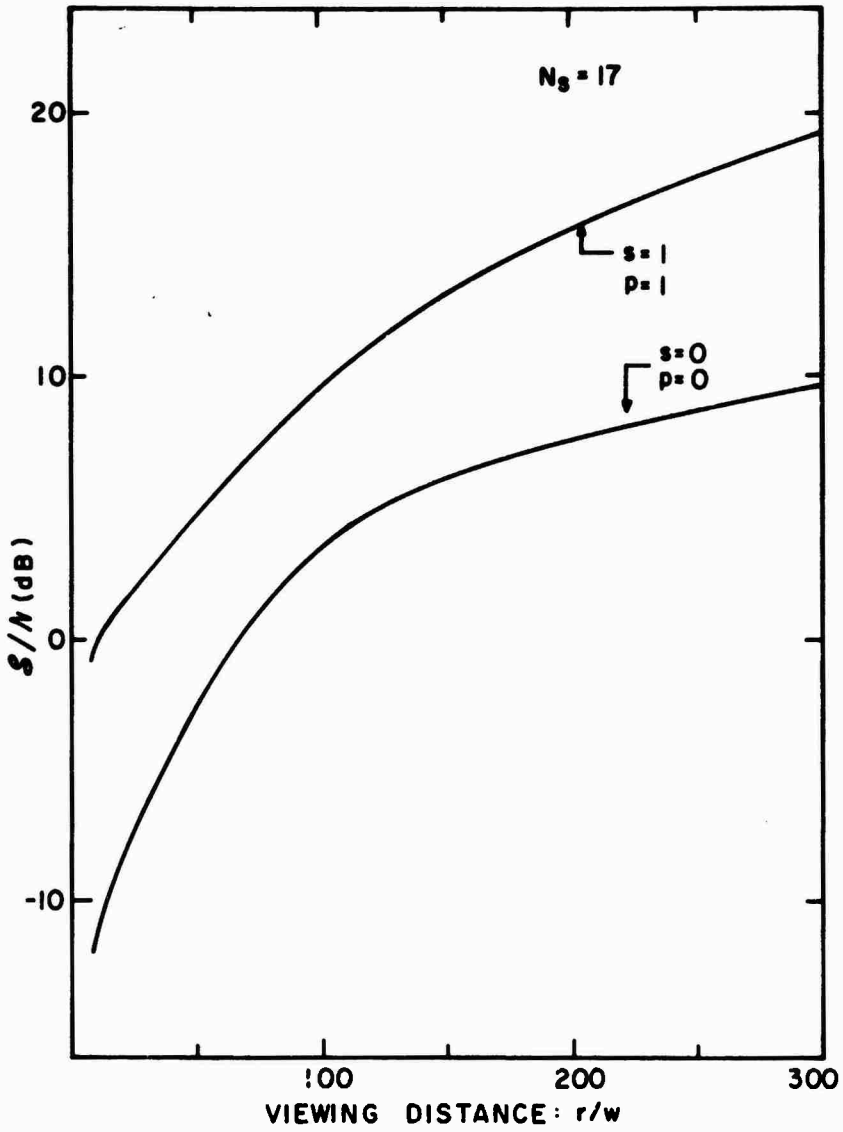


Figure 19. Perceived signal-to-noise ratio S/N as a function of viewing distance in picture widths for displays with N_S = 17 samples. The values of s and p are given in the figure.

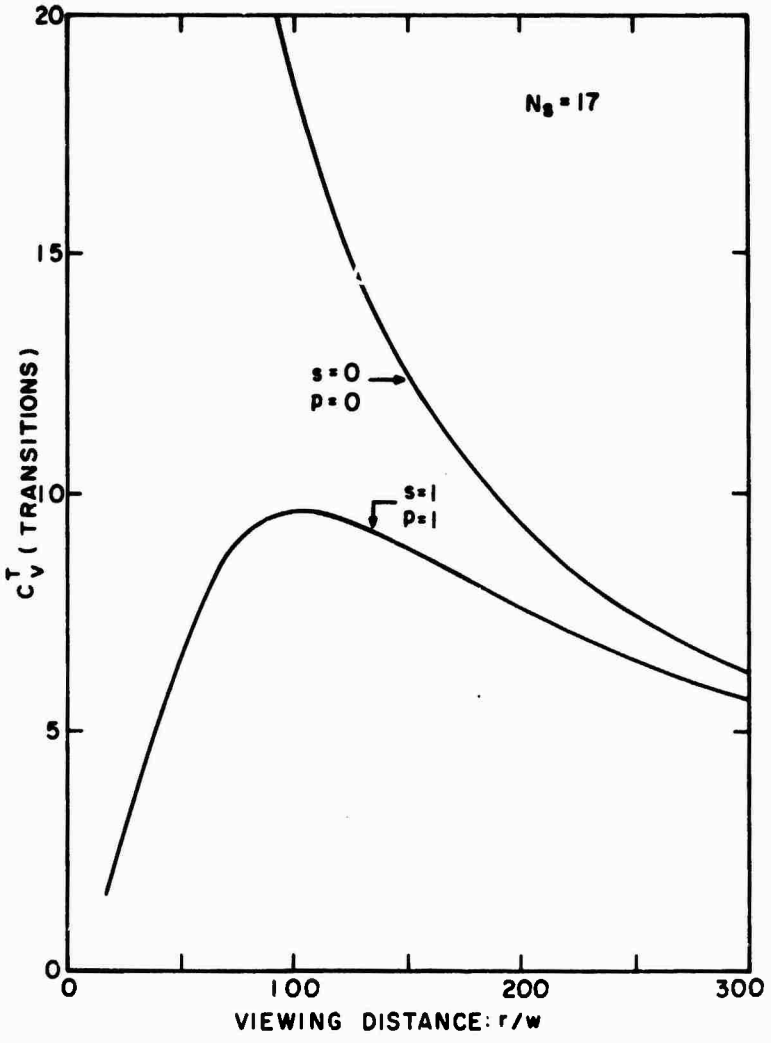


Figure 20. Visual capacity C_V^T as a function of viewing distance in picture widths for displays with N_S = 17 samples. The values of s and p are given in the figure.

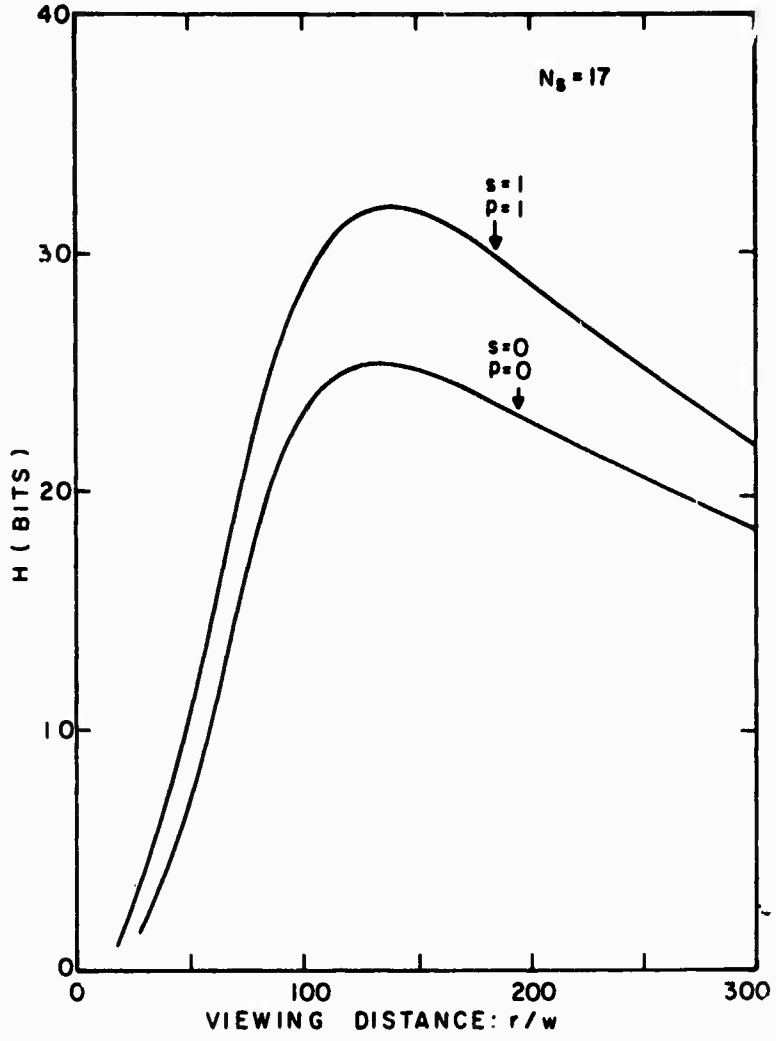


Figure 21. Total information capacity H as a function of viewing distance in picture widths for displays with N_s = 17 samples. The values of s and p are given in the figure. The value of the parameter $(S/N)_h$ was taken to be 15.

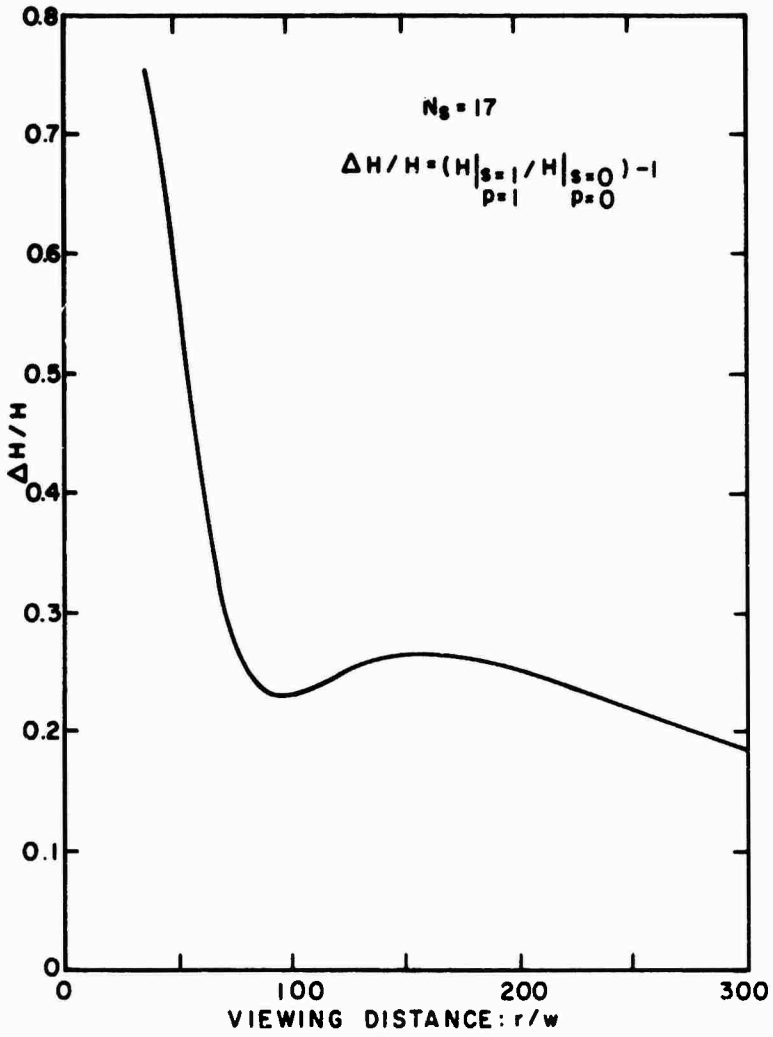


Figure 22. The fractional difference ΔH/H between the total information capacities of the displays of Fig. 21 as a function of viewing distance in picture widths.

viewing distance of 75 picture widths. The ordinate represents the integrands of Eqs. (50) and (51) for the perceived signal power S^2 and the perceived noise power N^2, respectively. The frequency coordinate is in units of cycles/degree-of-vision. A Lorentzian power specturm $\Phi(\omega) = B\,[\omega^2 + \omega_L^2]^{-1}$ was employed in all the calculations. The display sampling parameters used, s = p = 0, mean that we are considering delta-function sampling and printing, i.e., the input signal is sampled and printed over an infinitely narrow range within the sampling location. This is the arrangement that produces the maximum sampling noise. One sees from the figure that the signal part of the total spectrum is concentrated at the lowest retinal frequencies, rolling off rapidly above about 10 cycles/degree-of-vision. On the other hand, the noise power is concentrated in peaks centered around the sampling frequency (corresponding, at this viewing distance, to 22 cycles/degree-of-vision) and its harmonics and in a low-frequency peak near the value ν_c corresponding to the maximum sensitivity of the human visual system - 8 cycles/degree-of-vision.

To compute the perceived signal-to-noise ratio, according to Eqs. (50) and (51) we calculate the area under the perceived signal and noise power spectrum vs frequency curves, divide the signal area by the noise area, and take the square root. In Fig. 19 are plotted the perceived signal-to-noise ratios S(r)/N(r) so obtained as a function of viewing distance for the 17-element display with s = 0, p = 0 and for a display with s = 1 and a constant printing function (denoted by p = 1 in the figure) over one entire sampling location. The S(r)/N(r) for the display with s = 1, p = 1 rises rapidly with viewing distance, eventually approaching the predicted $(r/w)^2$ law, as predicted by Eq. (70). The S(r)/N(r) for the display with s = 0, p = 0 rises more slowly for large viewing distances, ultimately reaching a linear dependence on (r/w), as expected from the entry in Table 1 for undersampled displays with s = 0 and p = 0. The s = 1, p = 1 display has superior S/N at all viewing distances, the difference between the performance of the displays ranging between 6 and 10 dB over the range of r shown in Fig. 19. This behavior is consistent with our view that large sampling and printing widths favor signal-to-noise performance.

The visual capacity as a function of viewing distance is shown in Fig. 20 for the same two displays whose values of S/N are plotted in Fig. 19. The curve for the display with s = 0, p = 0 follows the simple 1/r dependence Eq. (28), since we have set $\hat{\Pi}(\omega) = R(\omega) = \mathrm{sinc}\,(s\omega/\omega_s) = 1$ in Eq. (27) for $C_v^T(r)$.

Recalling that the visual capacity represents perceived sharpness in the absence of noise, the results shown in Fig. 20 indicate that the display with s = 0, p = 0 should provide a greater impression of sharpness than the s = 1, p = 1 display. This is to be contrasted with the results of the S/N calculation, which indicated superior noise performance for the s = 1, p = 1 display.

In Fig. 21 is plotted the total information capacity H(r), obtained from the calculated $C_v^T(r)$ and S(r)/N(r) from Eq. (93). Here we have employed the value $(S/N)_h = 15$ for the effective signal-to-noise ratio of the human visual system. As can be seen from Fig. 21, H(r) for both displays has the same general form: a peak at a viewing distance of approximately 130 picture widths, a rapid fall-off at small viewing distances, and a more gradual drop at large viewing distances. At very large viewing distances, the curves will approach each other as the human visual system, not the display, becomes the limiting factor in the determination of the overall capacity of the display-observer channel. The curves of Fig. 21 indicate that the s = 1, p = 1 display provides, with the human observer, a higher overall information capacity at all viewing distances. This is better illustrated in Fig. 22, where the fractional difference $\Delta H/H$ between the s = 1, p = 1 and s = 0, p = 0 displays is shown as a function of viewing distance. Note the rapid rise of $\Delta H/H$ at small viewing distances. Figure 22 indicates that a simultaneous pair comparison of an s = 1, p = 1 display and an s = 0, p = 0 display would yield a threshold viewing distance, below which the s = 0, p = 0 display would appear far inferior to the s = 1, p = 1 display. This threshold viewing distance is calculated to be about 70 picture heights for illumination greater than about 100 ft-L. At lower brightnesses, the threshold will decrease markedly due to the shift of the peak of the MTF of the human visual system to smaller frequencies. For viewing distances between about 70 and 300 picture widths, $\Delta H/H$ is relatively slowly varying, averaging about 0.23. If we employ the value $(S/N)_h = 10$, instead of 15, $\Delta H/H$ in this range of viewing distances is reduced to an average of about 0.12.

f. Optimization of Displays. - Our method of analysis for one-dimensional sampled displays can be employed to optimize the performance of a display by seeking the maximum value of the total information capacity H as a function of the various sampling parameters (e.g., number of samples, sampling width,

printing function), the electronic bandwidth, and the viewing distance. External constraints, such as limitations on the viewing distance, number of samples, etc can also be accommodated.

In Fig. 23, we show the effect of varying the width of the printing function on the H(r) for the vertical (sampled) direction of television displays. The values of p, the width parameter for the Gaussian printing function Eq. (77), given in the figure, are the same as those employed to calculate the corresponding $C_v^T(r)$, shown in Fig. 9, and the S(r)/N(r), shown in Fig. 12. The value p = 0.642 was that measured for the real kinescope described in the discussion of the visual capacity in Section III.D. Comparison of Figs. 9, 12, and 23 shows that, whereas small values of p result in enhanced visual capacity, and large values of p favor signal-to-noise performance, the behavior of H(r) with p is substantially more complicated. The value of p for optimum H(r) clearly depends on viewing distance. At small viewing distances $r/h \lesssim 3$, H(r) is enhanced by increasing the printing width p from its observed value, indicating that H(r) is dominated by signal-to-noise considerations. On the other hand, at larger viewing distances, H(r) is enhanced by employing smaller values of p, showing that the value of S(r)/N(r) is sufficiently large to be able to support some degradation in exchange for increased sharpness. The absolute maximum value of $H(r_{opt})$ is achieved for a value of p approximately 20% smaller than the observed value. Decreasing p beyond this point reduces $H(r_{opt})$ but gives some advantage at very large viewing distances. The value of r_{opt} itself achieves a minimum value of approximately 5 picture heights. Values of p that are too small or too large act to increase the optimum viewing distance. As a general comment, we call attention to the shape of the curves; the curve of H(r) with the largest value of p, having the best noise performance but the worst "sharpness," has a soft, gradual rise at small r and a comparatively broad peak, whereas the curve with the smallest value of p, having the worst noise performance but the best "sharpness," has an abrupt rise and a sharper peak. The abrupt increase occurs after the value of S(r)/N(r) passes through the value unity [see Fig. 12].

At any particular viewing distance, we can compute the optimum value of the printing width, subject to the constraint that all other display parameters are held constant. An example of such an optimization technique is shown in

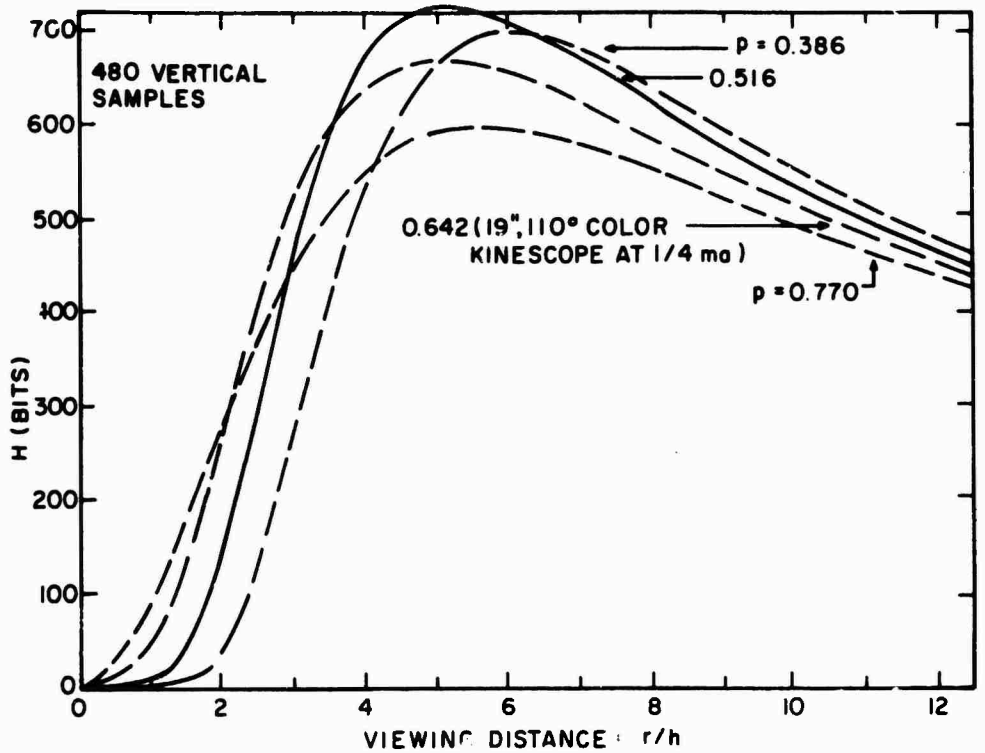


Figure 23. Total information capacity H as a function of viewing distance in picture heights for the vertical direction (N_s = 480) of a television display. The values of the parameter p, which represents the effective width of the Gaussian printing function, are given in the figure. The particular value obtained from measurements of a commercially available kinescope is also indicated. The parameter $(S/N)_h$ was taken to be 10.

Fig. 24. Consider a hypothetical 500-sample display with no band limitation applied to the input signal. Let us assume that technical reasons have forced us to consider only delta-function sampling (s = 0) and printing functions whose intensity profile is constant over a fractional width p of each sampling aperture on the display screen and is zero everywhere else [$\hat{\Pi}(\omega)$ = sinc $(p\omega/\omega_s)$]. We wish to know the optimum value of p for various viewing distances. Figure 24 plots the calculated H as a function of s (taking $(S/N)_h$ = 15) for several viewing distances. From the figure, it is seen that, as the viewing distance increases from 2 to 14 picture widths, the optimum value of p decreases from unity to the value 0.56. Furthermore, as the viewing distance increases, the sensitivity of H to the specific value of p decreases. For example, if p is decreased from its optimum value by 0 1, the cost would be a 46% drop in H at a viewing distance of 2 picture widths, but only 2% at 5 picture widths. This

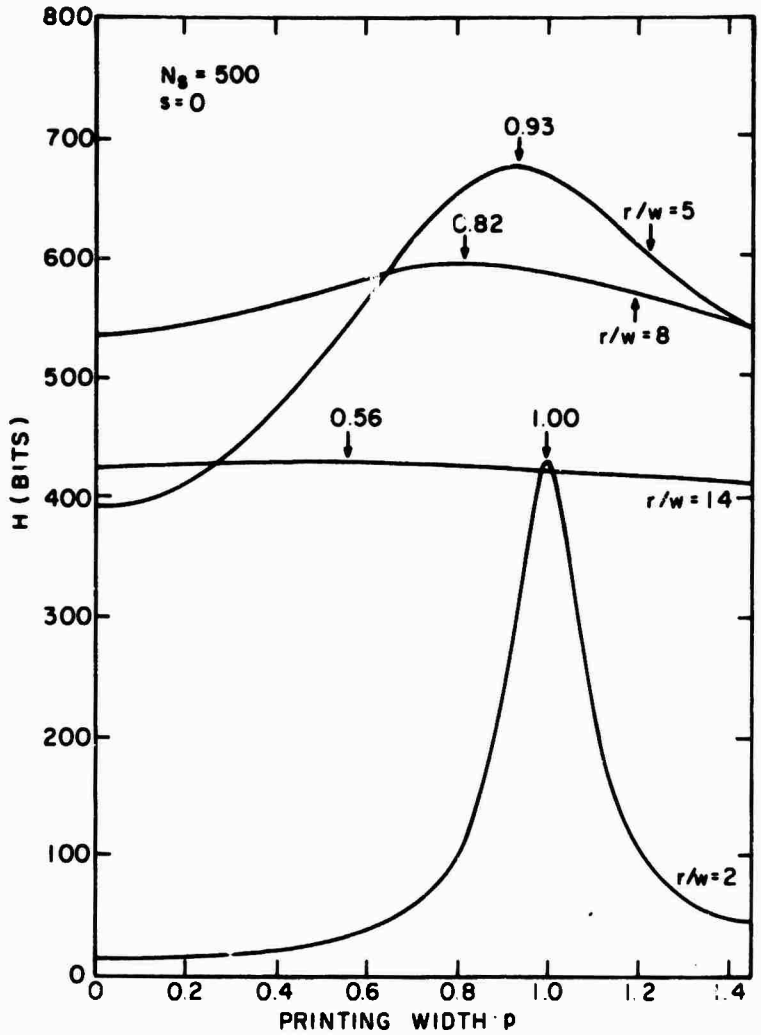


Figure 24. Total information capacity H as a function of effective printing width p for a hypothetical 500-sample display with s = 0 and $\omega_M = \infty$. Here p is the fractional width of each sampling location that is activated with a constant intensity. The values of the viewing distance, expressed in picture widths, are given in the figure. The vertical arrows indicate the values of p for maximum H. The parameter $(S/N)_h$ was taken to be 15.

behavior can be understood as follows. At small viewing distances, the display-observer system is optimized by full-width printing because of its superior signal-to-noise characteristics. As the viewing distance increases, the perceived signal-to-noise ratio for p = 1 becomes large enough so that sampling noise is far less influential in limiting total information capacity. It then pays to decrease p in order to gain the larger perceived sharpness that is

characteristic of small values of p. However, as the viewing distance continues to increase, the human visual system becomes the limiting factor in determining H; it then makes less and less difference just how the display is designed (for a given number of samples). Thus, curves of H vs p become progressively flatter.

If we relax the constraint s = 0 and allow s to take on any value, we have a maximization problem in two dimensions; at any viewing distance, H must be maximized with respect to both p and s. For our hypothetical 500-element display, at a viewing distance of 2 picture widths, it turns out that p = 1 gives a relative maximum value of H regardless of the value of s. Thus, Fig. 25, which graphs H against s for p = 1, respresents the locus of relative maximum values of H as the sampling width is varied. The absolute maximum value of H is achieved for the combination of parameter s = 0.43, p = 1. This value of H is 11% larger than the value for s = 0, p = 1, which is the maximum achievable H subject to the constraint s = 0. In a similar manner, one can perform the two-dimensional maximization calculation at other viewing distances. For example, we calculate that, in increasing the viewing distance from 2 to 5 picture widths, the optimum pair of s, p parameters shifts from s = 0.43, p = 1 to s = 0.99, p = 0.99.

The above example by no means exhausts the variables of the optimization problem. We may introduce an electronic filter function of finite passband before the sampling process is performed. The width of the passband then becomes a third variable which must be optimized simultaneously with s and p. In addition, we may, of course, allow other forms of printing functions in addition to the simple constant P(x) of fractional width p.

The results of any optimization calculation, once performed at a particular viewing distance for a given number of samples, cannot be assumed to be valid if the number of samples is changed. Figure 26 shows curves of H as a function of the number of samples for 3 different sampling widths with the viewing distance held constant at 5 picture widths and with a constant printing function width p = 1. All the curves have the same general form, rising rapidly at first but eventually saturating at a value determined by the limitations of human visual system. It is noteworthy that, whereas full-width sampling s = 1 is favored for greater than about 300 samples, half-width sampling s = 1/2 gives a larger value of H when the number of samples is less than 300. Indeed,

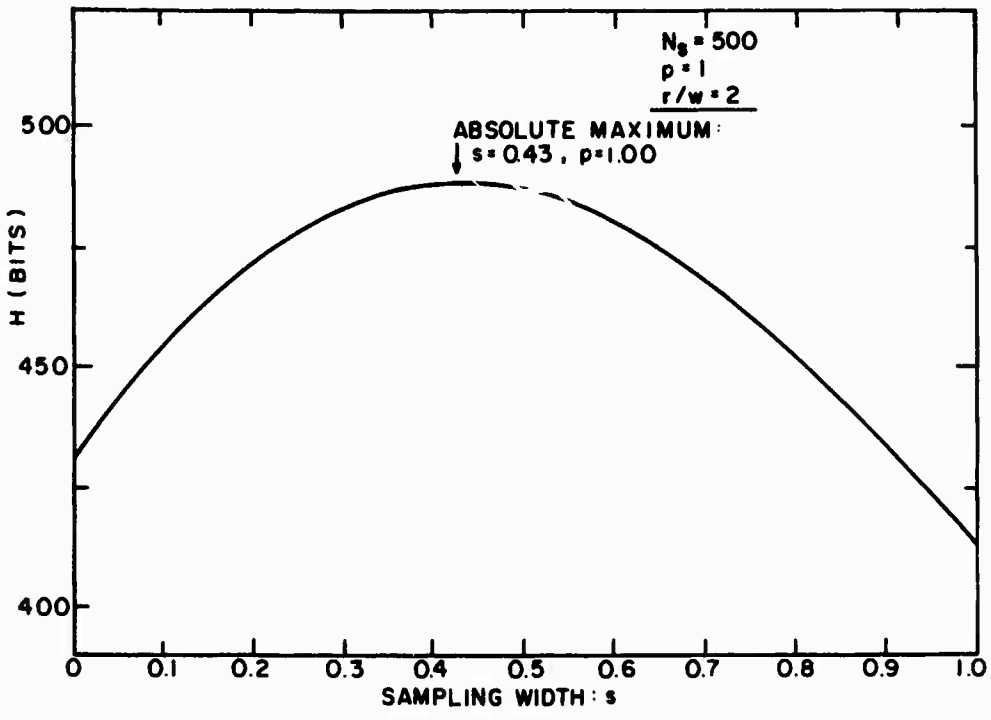


Figure 25. Total information capacity H as a function of sampling width s for a hypothetical 500-sample display with p = 1 and $\omega_M = \infty$. The viewing distance is 2 picture widths. The curve represents the locus of relative maximum values of H in the two-dimensional space of s, p values. The absolute maximum of H is indicated in the figure. The parameter $(S/N)_h$ was taken to be 15.

when the number of samples is less than approximately 225, even s = 0 sampling gives a larger value of H than does the case s = 1. This example serves to emphasize that the number of display samples must be explicitly included in any optimization calculation.

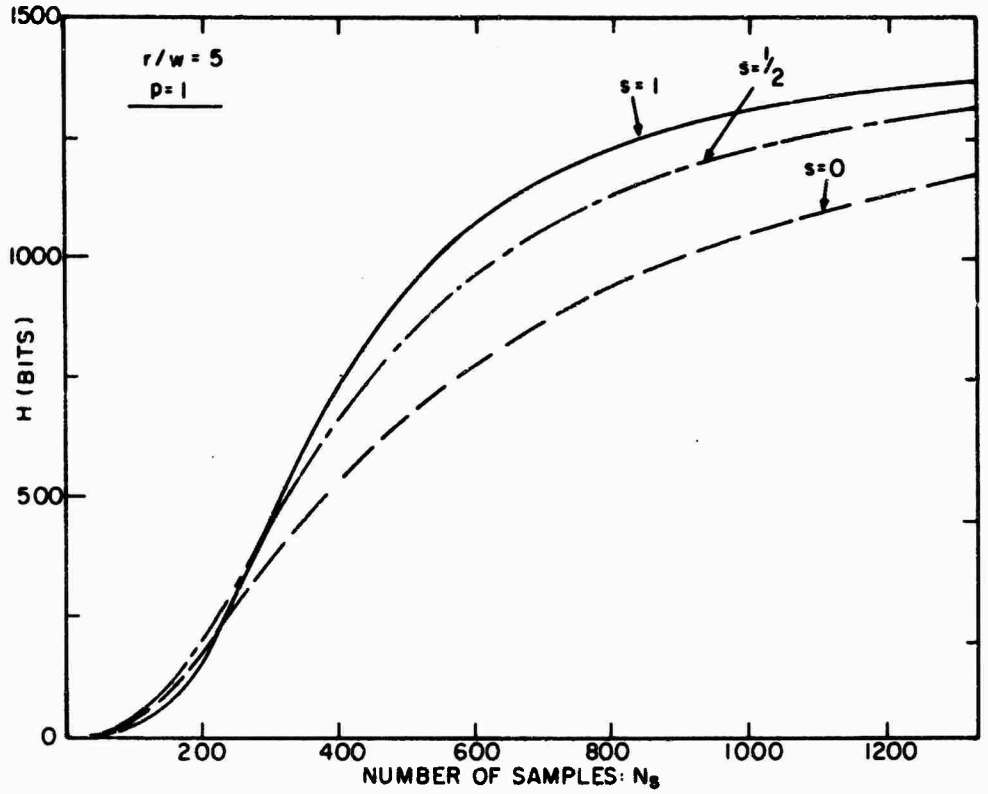


Figure 26. Total information capacity H as a function of the number of samples N_s for hypothetical displays with p = 1 and $\omega_M = \infty$. The viewing distance is 5 picture widths. The values of s are given in the figure. The parameter $(S/N)_h$ was taken to be 15.

SECTION IV

TWO-DIMENSIONAL ANALOG DISPLAYS - THE VISUAL CAPACITY

A. MATHEMATICAL FORMULATION

The visual capacity [11] for one-dimensional analog displays has been defined in terms of the perceived visual response to sharp edge inputs. The calculated edge width x_e, as perceived by the observer, then determines the maximum number of resolvable edge transitions across a display of width w through the simple formula $C_v^T = w/x_e$. As discussed in ref. [11] and Section III.D of this report, C_v^T represents the information capacity of a noiseless two-level communication system. For noisy multilevel displays, the total information capacity H is proportional to C_v^T [Eq. (93)]. In the one-dimensional case, the analogy between visual capacity and the maximum pulse transmission rate μ of a communication channel is straightforward. For the case of two-dimensional displays, no direct analogy exists for the simple reason that there is only one time dimension. Nevertheless, if we continue to define information capacity as the base-two logarithm of the total number of discrete states the system can assume, it is clear that the information capacity of a two-dimensional display will be proportional to the number of discrete locations that can be perceived within a display of area A. Accordingly, we define the two-dimensional visual capacity C_{v2}^T as the number of perceivable *spots* within the area A:

$$C_{v2}^T = A/\sigma_e \tag{109}$$

Here σ_e is the perceived area of a single spot - the response of the display-observer system to a delta-function point input.

We shall calculate an expression for σ_e in a manner analogous to the calculation of the quantity x_e for the one-dimensional case. Let distance on the display screen be described by the polar coordinates ρ and α, where α is the polar angle, as measured from the horizontal, and ρ is the radial distance from a point on the screen to the origin of the coordinate system. The location of the origin is arbitrary. We define the psychophysical response function $E_\delta(\rho,\alpha)$ as the perceived response to a unit delta-function input signal. The meaning of the unit delta-function is as follows. The brightness units are defined such that the integral over the display screen of the response of the display to a

delta-function input of strength D is unity. Denoting the intensity pattern on the screen by $I_\delta(\rho,\alpha)$ and taking the delta-function input at the origin, we must have

$$\int_0^\infty \rho\, d\rho \int_0^{2\pi} d\alpha\, I_\delta(\rho,\alpha) = 1 \tag{110}$$

where

$$I_\delta(\rho,\alpha) = D \int_0^\infty \frac{\omega\, d\omega}{4\pi^2} \int_0^{2\pi} d\phi\, R_o(\omega,\phi)\, \exp[i\omega\rho\, \cos\,(\phi-\alpha)] \tag{111}$$

Here $R_o(\omega,\phi)$ is the overall two-dimensional MTF of the display system as a function of the magnitude ω and direction ϕ (with respect to the horizontal) of the two-dimensional spatial frequency vector. The response $I_\delta(\rho,\alpha)$ is assumed to be sufficiently localized for the integration over ρ in Eq. (110) to be extended to infinity. Substituting Eq. (111) into Eq. (110) and performing the spatial integration, it is a straightforward exercise in Fourier analysis to show that our definition of a unit delta-function implies simply that its strength is unity:

$$D = 1 \tag{112}$$

To obtain Eq. (112), we have employed the identity $R(0,\phi) = 1$, the condition that there be no amplification or attenuation of a dc input.

In the spirit of linear systems analysis, as employed for the case of the one-dimensional visual capacity, we shall compute the perceived response $E_\delta(\rho,\alpha)$ using a two-dimensional MTF $O(\nu,\phi)$ to describe the processing performed by the human visual system. Here $O(\nu,\phi)$ describes the perceived contrast for a sinusoidal grating of retinal frequency $\nu = \omega r/2\pi$, oriented at an angle ϕ from the horizontal.* Then, just as in the one-dimensional case, the perceived response $E_\delta(\rho,\alpha)$ is obtained from Eq. (111) by simply multiplying the integrand of this equation by the function $O(\omega r/2\pi,\phi)$. Setting $D = 1$, we have

$$E_\delta(\rho,\alpha) = \int_0^\infty \frac{\omega d\omega}{4\pi^2} \int_0^{2\pi} d\phi\, O(\omega r/2\pi,\phi)\, R_o(\omega,\phi)\, \exp[i\omega\rho\, \cos\,(\phi-\alpha)] \tag{113}$$

*This is an important assumption that must be justified by experiment. See Section VI for initial experimental results that indicate that such an assumption is valid for complex two-dimensional gratings.

We note that, with the normalization condition Eq. (110), both $I_\delta(\rho,\alpha)$ and $E_\delta(\rho,\alpha)$ have the dimensions of inverse area. In fact, the magnitude of E_δ, say near $\rho = 0$, is on the order of $1/\sigma_e$, where σ_e can be thought of as the inverse square of an effective cutoff frequency for the combined display-observer system. (For example, a "perfect" system would have unity response at all frequencies, so that $\sigma_e = 0$. Then E_σ diverges, as it should for a perfectly reproduced delta-function.) The effective display area covered by E_σ is on the order of σ_e. This can be understood by noticing that the range of ρ in the integral of Eq. (113) is determined by the fact that contributions for large arguments of the exponential in Eq. (113) tend to oscillate rapidly. These oscillations drastically reduce contributions to the integral if $|\omega\rho \cos(\phi-\alpha)| \gtrsim \pi$, i.e., if the value of ρ, averaged over all angles, is much greater than $\sigma_e^{1/2}$, the inverse effective cutoff frequency for the display-observer system. Thus, a simple, mathematically convenient definition of the perceived area of a single spot σ_e is in terms of the integral,

$$1/\sigma_e = \int_0^\infty \rho \, d\rho \int_0^{2\pi} d\alpha \; E_\delta^2(\rho,\alpha) \tag{114}$$

Equation (114) defines the perceived spot area σ_e as the area integral of the square of the effective overall point spread function for a unit delta-function input. It is directly analogous to the geometrical definition of the perceived width of a single edge, employed in ref. [11] in the derivation of the one-dimensional visual capacity.

Equation (114) for σ_e can be expressed in terms of the display and observer MTF's by substituting Eq. (113) for $E_\delta(\rho,\alpha)$ into Eq. (114). Performing the spatial integration gives rise to a two-dimensional delta-function in frequency, which allows one of the two-dimensional frequency integrations to be performed trivially. We are left with

$$1/\sigma_e = \int_0^\infty \frac{\omega d\omega}{4\pi^2} \int_0^{2\pi} d\phi \; O(\omega r/2\pi,\phi) \; O(\omega r/2\pi, \phi + \pi)$$

$$\times R_o(\omega,\phi) \; R_o(\omega,\phi + \pi) \tag{115}$$

Next, from Eq. (113), it is easy to show that the reality of $E_\delta(\rho,\alpha)$ requires that

$$R_o(\omega,\phi + \pi) = R_o^*(\omega,\phi) \tag{116a}$$

$$O(\omega,\phi + \pi) = O^*(\omega,\phi) = O(\omega,\phi) \tag{116b}$$

Using the results Eqs. (115) and (116), Eq. (109) gives us the expression for the visual capacity,

$$C_{v2}^T(r) = A \int_0^\infty \frac{\omega d\omega}{4\pi^2} \int_0^{2\pi} d\phi \;\; O^2(\omega r/2\pi,\phi) \;\; |R_o(\omega,\phi)|^2 \tag{117}$$

Comparing Eq. (117) with Eq. (26) for the one-dimensional visual capacity, we see that both expressions involve the integral of the square of the magnitude of the combined MTF of the display and the human visual system. Equation (117) explicitly includes the effect of any anisotropy of the display or observer response. It should be emphasized that C_{v2}^T represents the number of perceivable *spots* within a display of area A, whereas C_v^T represents the number of perceivable *edges** across a display of width w. Accordingly, we expect that C_{v2}^T is a much larger number than C_v^T. Indeed, one might expect that, for nearly isotropic displays, $C_{v2}^T \cong (C_v^T)^2$. In the following, we shall see to what extent this expectation is valid.

*The one-dimensional visual capacity was defined in terms of x_e, the effective edge transition width. However, it can be shown rigorously that x_e also represents the effective width of the response of the display-observer system to a delta-function input. Therefore, C_v^T also represents the number of perceivable line segments across a one-dimensional display of width w.

B. PROPERTIES OF $C_{v2}^{T}(r)$

1. Far-Field Viewing

In this section, we consider sufficiently large viewing distances such that the following inequality is satisfied:

$$r/2\pi\nu_o \gg 1/\text{Min}\,[\omega_M(\phi)] \tag{118}$$

where $\text{Min}[\omega_M(\phi)]$ represents the minimum value of the cutoff frequency of the display MTF $R_o(\omega,\phi)$. In that case, $R_o(\omega,\phi)$ in the integrand of Eq. (117) may be replaced by unity, so that $C_{v2}^{T}(r)$ is given by

$$C_{v2}^{T}(r) = A \int_0^{\infty} \frac{\omega d\omega}{4\pi^2} \int_0^{2\pi} d\phi\; O^2(\omega r/2\pi,\phi) \tag{119}$$

We now approximate $O(\omega r/2\pi,\phi)$ by the isotropic MTF $O(\nu)$. This approximation is supported by the experimental results presented elsewhere in this report (see Section VI). In that case Eq. (119) becomes, after transforming to the retinal frequency coordinate $\nu = \omega r/2\pi$,

$$C_{v2}^{T}(r) = (2\pi A/r^2) \int_0^{\infty} d\nu\; \nu\; O^2(\nu) \tag{120}$$

From the measured $O(\nu)$ (Fig. 5), we compute numerically

$$\int_0^{\infty} d\nu\; \nu\; O^2(\nu) = 2.57\; \nu_o^{\,2} = 0.154/(\theta_e^{\infty})^2 \tag{121}$$

where we have employed the value $\theta_e^{\infty} = 1.84$ min of angle [Eq. (30)] for the

perceived angular width of a single perfectly reproduced edge. From Eqs. (120) and (121), the expression for $C_{v2}^{T}(r)$ at large viewing distances is

$$C_{v2}^{T} = 0.970\ A/(\theta_e^{\infty} r)^2 \tag{122}$$

valid when the inequality Eq. (118) is met. The physical meaning of Eq. (122) is that, at large viewing distances, the two-dimensional visual capacity is simply the solid angle A/r^2 subtended by the display divided by $1.031(\theta_e^{\infty})^2$, the perceived solid angle subtended by a single perfectly reproduced delta-function spot. Comparing Eq. (122) with the expression [Eq. (28)] $C_v^T(r) = w/\theta_e^{\infty} r$ for the one-dimensional visual capacity at large viewing distances, we see that, for a rectangular display, C_{v2}^{T} is nearly exactly equal to the product of the one-dimensional visual capacities for the horizontal and vertical directions. The small numerical difference is geometrical in origin and arises from the definitions of x_e and σ_e.

2. Near-Field Viewing

We now consider sufficiently small viewing distances such that the display cutoff frequency $\omega_M(\phi)$ always corresponds to retinal frequencies much smaller than ν_o:

$$r/2\pi\nu_o << 1/\text{Max}\,[\omega_M(\phi)] \tag{123}$$

Thus, we can replace $O(\nu,\phi)$ in Eq. (117) by the low-frequency linear asymptote Eq. (32). Equation (117) becomes

$$C_{v2}^{T}(r) = 0.335\ A\ (r/2\pi\nu_o)^2 \int_0^{2\pi} d\phi \int_0^{\infty} d\omega\ \omega^3\ |R_o(\omega,\phi)|^2 \tag{124}$$

From Eq. (124), it is seen that, in near-field viewing, $C_{v2}^{T}(r)$ increases as the square of the viewing distance, as does the one-dimensional visual capacity [Eq. (33)]. Thus, even for isotropic displays, $C_{v2}^{T}(r)$ is not proportional to $[C_v^T(r)]^2$ at small viewing distances. This result is profound and requires special comment. It arises from our representation of the two-dimensional response of the human visual system by a function $O(\nu,\phi)$, defined

in terms of the perceived response to a one-dimensional sine-wave grating with retinal frequency of magnitude ν, oriented at an angle ϕ from the horizontal. This statement, along with the use of the low-frequency linear asymptote derived from one-dimensional sine-wave measurements [6] at $\phi = \pi/2$, leads directly to the result that both C_v^T and C_{v2}^T are proportional to r^2 at small viewing distance. It is possible to construct a model in which this is not the case. Suppose, for example, the visual system operated with independent, highly orientation-specific frequency sensors, sensitive to the component of the vector $\vec{\nu}$ along two mutually perpendicular directions, say $\phi = 0$ and $\phi = \pi/2$. In that case, we might expect $O(\omega r/2\pi, \phi)$ in Eq. (117) to be a separable function of the components of $\vec{\nu}$ along $\phi = 0$ and $\phi = \pi/2$, i.e., $O(\omega r \cos \phi/2\pi)\ O(\omega r \sin \phi/2\pi)$, where $O(\nu)$ represents the measured [6] one-dimensional MTF. Then, Eq. (117) would indeed yield $C_{v2}^T(r) \sim [C_v^T(r)]^2$ for isotropic displays at small viewing distances. However, experimental evidence, presented in Section VI, supports the existence of a nearly isotropic $O(\nu, \phi)$.

As an example of the effect of display anisotropy on $C_{v2}^T(r)$ at small viewing distances, we consider a hypothetical anisotropic display with elliptical symmetry. The passband is flat [$R_o(\omega, \phi) = 1$] with a maximum frequency ω_{Mh} in the horizontal direction and ω_{Mv} in the vertical direction. The principal axes of the elliptically shaped two-dimensional passband are assumed to coincide with the horizontal and vertical directions. We define ω_{Mo} as the geometric mean of ω_{Mh} and ω_{Mv}:

$$\omega_{Mo} = (\omega_{Mh}\ \omega_{Mv})^{1/2} \tag{125}$$

The dimensionless anisotropy factor $\underline{a}$ is defined according to

$$a = \frac{\omega_{Mv} - \omega_{Mh}}{\omega_{Mv} + \omega_{Mh}} \tag{126}$$

The range of $\underline{a}$ is $-1 \leq a \leq +1$, with $a = 0$ corresponding to the circular passband of an isotropic display. With the definitions Eqs. (125) and (126), the cutoff frequency $\omega_M(\phi)$ for elliptical symmetry is given by

$$\omega_M(\phi) = \omega_{Mo}\left[\frac{1 - a^2}{1 + 2a\cos 2\phi + a^2}\right]^{1/2} \tag{127}$$

Then Eq. (124) for $C_{v2}^T(r)$ in the near-field viewing limit becomes

$$C_{v2}^T(r) = 0.335\ A\ (r/2\pi\nu_o)^2 \int_0^{\pi/2} d\phi\ \omega_M^4(\phi) \tag{128}$$

The integration over angle can be performed analytically, giving us the result,

$$C_{v2}^T(r) = 0.526\ A\ \omega_{Mo}^4\ (r/2\pi\nu_o)^2\ [(1 + a^2)/(1 - a^2)];$$
$$\text{for } r/2\pi\nu_o << \omega_{Mo}^{-1}\ [(1 - |a|)/(1 + |a|)] \tag{129}$$

For a constant passband area $\pi\omega_{Mo}^2$, the effect of the anisotropy is entirely contained in the factor $(1 + a^2)/(1 - a^2)$. We see that $C_{v2}^T(r)$ is enhanced by display anisotropy at small viewing distances. This is a direct result of the quadratic increase of $O^2(\nu)$ at low retinal frequencies, making it profitable to trade off increased bandwidth over a certain range of angles against decreased bandwidth at other angles.

3. Maximum Visual Capacity

Just as in the case of the one-dimensional visual capacity, $C_{v2}^T(r)$ exhibits a maximum at a viewing distance r_p, which depends on the band-limiting characteristics of the display system. As a simple example, we once agai. take the case of a flat passband with elliptical symmetry $R(\omega,\phi) = 1$ for $\omega \leq \omega_M(\phi)$, with $\omega_M(\phi)$ given by Eq. (127). Differentiating Eq. (117) for C_{v2}^T with respect to r, the condition for maximum visual capacity is

$$\int_0^{2\pi} d\phi \int_0^{\omega_M(\phi)r_p/2\pi} d\nu\ \nu^2\ \frac{d}{d\nu}\ O^2(\nu) = 0 \tag{130}$$

After a partial integration of Eq. (130), the maximum visual capacity can be written in the form

$$C_{v2}^{T}(r_p) = (A/8\pi^2) \int_0^{2\pi} d\phi\ \omega_M^2(\phi)\ O^2[\omega_M(\phi) r_p/2\pi] \tag{131}$$

For an isotropic display a = 0, Eq. (130) can be easily solved numerically, using the measured [6] $O(\nu)$, to give

$$\begin{aligned} r_p &= 1.32\ (2\pi\nu_o/\omega_{Mo}) \\ &= 3.80 \times 10^3/\omega_{Mo} \end{aligned} \tag{132}$$

From Eq. (131), the maximum visual capacity is

$$\begin{aligned} C_{v2}^{T}(r_p) &= \frac{\pi}{4}\ (A\ \omega_{Mo}^2/\pi^2)\ O^2(1.32\ \nu_o) \\ &= 1.00\ (A\ \omega_{Mo}^2/\pi^2) \end{aligned} \tag{133}$$

where we have used the value $O(1.32\ \nu_o) = 1.13$ (Fig. 5). From Eqs. (132) and (133), we can make two interesting observations. First, comparing the viewing distance r_p, given in Eq. (132), with the result [11] for the equivalent one-dimensional display shows that r_p is about 11% smaller for C_{v2}^{T} than for C_v^T. This difference is not considered significant. Second, since the quantity $A\omega_{Mo}^2/\pi^2$ represents the product of the number of TV lines for the horizontal and vertical directions of a rectangular display, $C_{v2}^{T}(r_p)$ is numerically equal to the product of the maximum values of the one-dimensional visual capacities for the horizontal and vertical directions.

For the case of anisotropic displays $|a| > 0$, Eqs. (130) and (131) must be solved using a computer. However, we can determine the manner in which anisotropy affects r_p and $C_v^T(r_p)$ by performing a perturbation calculation. If $|a| \ll 1$, we can expand the integral in Eq. (130) in a Taylor series in $\underline{a}$. The procedure is straightforward but somewhat lengthy. Equation (127) for $\omega_M(\phi)$ is expanded in a Taylor series, keeping terms on the order of a^2. The result is inserted into Eq. (130), and the first two derivatives of the integral with respect to $\underline{a}$ are calculated. We permit r_p

to vary from its a = 0 value in order to ensure that the integral vanishes. We find that the first non-vanishing contribution to the change δr_p in the viewing distance for maximum C^T_{v2} is on the order of a^2:

$$\delta r_p / r_p = -\frac{1}{4} a^2 \left\{ 1 + \left[\frac{d \ln (d\, O^2(\nu)/d\nu)}{d \ln \nu} \right] \right\} \Bigg|_{\nu = 1.32\, \nu_o} \tag{134}$$

From the measured $O(\nu)$, we estimate the quantity in the square brackets to be 1.3. Thus, Eq. (134) becomes

$$\delta r_p / r_p = -0.57\, a^2; \text{ for } |a| \ll 1 \tag{135}$$

In a similar manner, we find numerically,

$$\delta C^T_{v2}(r_p) / C^T_{v2}(r_p) = -0.75\, a^2; \text{ for } |a| \ll 1 \tag{136}$$

where $\delta C^T_{v2}(r_p)$ is the change in the maximum value of C^T_{v2} due to anisotropy.

From Eqs. (135) and (136), it is seen that the effect of anisotropy is to lower the maximum two-dimensional visual capacity and shift the position of the peak to smaller viewing distances. Because of the a^2 dependence, the effects are actually quite small for reasonable values of $\underline{a}$. From Eq. (136), we see that the value $|a| \gtrsim 0.37$ is needed to lower $C^T_{v2}(r_p)$ by 10% from its a = 0 value. This corresponds to an aspect ratio $\omega_{Mh}/\omega_{Mv} \gtrsim 2.2$ [see Eq. (126)]. Even larger aspect ratios are required to reduce r_p by 10%. These results are consistent with early experimental work [22] which found that, for spot aspect ratios less than about 2:1, the resolution limit of a spot is determined only by its area.

C. A CALCULATED EXAMPLE

As an illustrative example of the above results, we have calculated $C^T_{v2}(r)$ for a hypothetical anisotropic display. We assume that the display is characterized by a two-dimensional MTF with elliptical symmetry and that the passband is flat, with the bandwidth $\omega_M(\phi)$ determined by Eq. (128). In

the calculation, the area of the passband $\pi\omega_{Mo}^2$ was held constant as $\underline{a}$ was varied in order to isolate the effect of anisotropy. We assumed a square display of width w, so that $A = w^2$. The results are shown in Fig. 27, where C_{v2}^T is plotted as a function of viewing distance for a particular value of the area of the passband corresponding to a $N_{TV} = \omega_{Mo}w/\pi = 366.6$ line display for $\underline{a} = 0$. Curves are shown for $\underline{a} = 0$, 1/3, 2/3, 5/6, and 0.98, corresponding to values of the frequency ratio $\omega_{Mv}/\omega_{Mh} = 1$, 2, 5, 11, and 99, respectively. From the figure we note the following important conclusions:

(1) Large anisotropies are required in order to produce an appreciable effect; for $\underline{a} = 1/3$, the drop in the peak of C_{v2}^T from its $\underline{a} = 0$ value is only about 8% [in agreement with Eq. (136)];

(2) As the anisotropy is increased, the peak of C_{v2}^T is depressed and moves to smaller viewing distances [in agreement with Eq. (135)];

(3) For large anisotropies, C_{v2}^T is enhanced at very small viewing distances [in agreement with Eq. (129)], whereas it is strongly depressed at intermediate viewing distances.

D. SUMMARY

Our results indicate that for noiseless, analog displays, the one-dimensional visual capacity C_v^T is adequate for treating two-dimensional displays whose anisotropy factor $|a|$ is less than about 1/3 (frequency ratios $\lesssim 2$). This conclusion is based on the small effect of anisotropy on C_{v2}^T for $|a| \lesssim 1/3$ and on our derived results showing that C_{v2}^T is numerically almost exactly equal to the product of the one-dimensional visual capacities for the horizontal and vertical directions at their respective maxima and at very large viewing distances. This is a gratifying result, for it shows that the simpler one-dimensional descriptor can be applied to most practical situations. For highly anisotropic displays, the two-dimensional visual capacity gives results that cannot be simply obtained from the one-dimensional counterpart. In these cases Eq. (117) for C_{v2}^T can be employed to calculate the effect of anisotropy on the number of perceivable spots at any viewing distance.

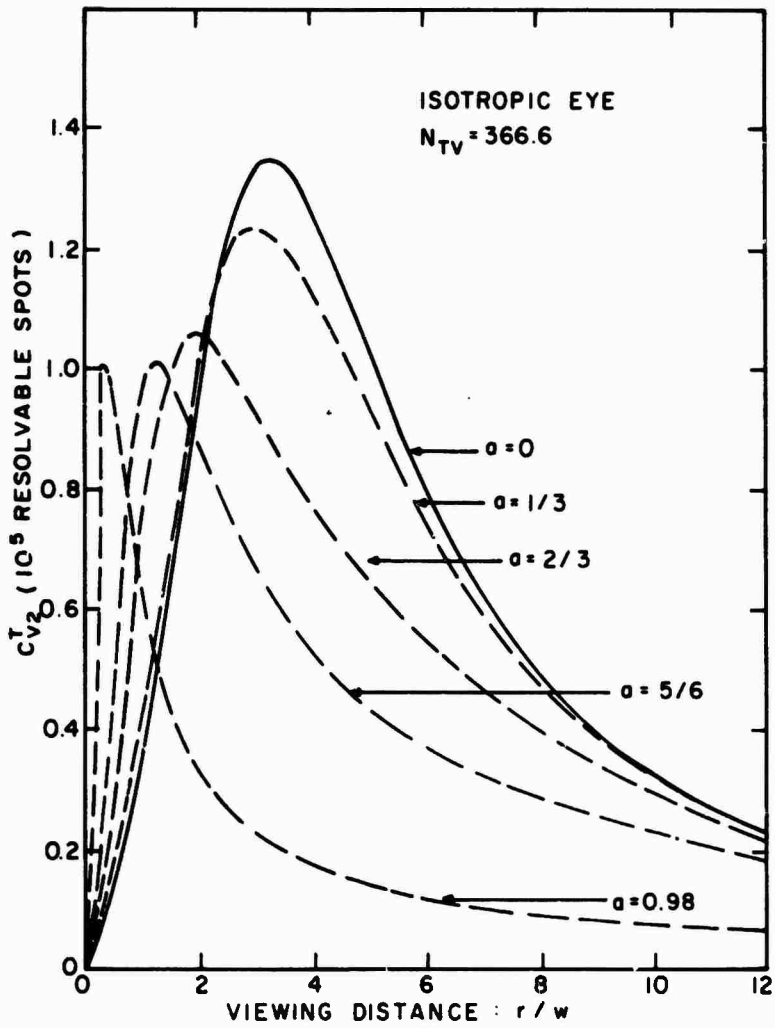


Figure 27. Two-dimensional visual capacity C^T_{v2} as a function of viewing distance in picture widths for a hypothetical anisotropic analog display. The display is characterized by a flat two-dimensional MTF with elliptical symmetry. The area of the passband corresponds to that for an isotropic display with N_{TV} = 366.6. The values of the anisotropy factor a are given in the figure.

SECTION V

OPTICAL BLOCK PROCESSOR

To provide experimental input to the analytical studies described in this report, a real-time optical block processor was constructed that has the capability of producing sampled pictures with a wide variety of sampling and printing functions. The basic technique utilized an array of plexiglass "light pipes" to subdivide and average the luminance of an image into a finite number of picture elements. The processor has the capability of handling, in a simple and direct way, both color and black and white images, without the expense and complication of digital processing technicques.

A sketch of the apparatus is shown in Fig. 28, and a photograph of the actual device constructed is shown in Fig. 29. The apparatus consists of 1050 uniform size plexiglass blocks (0.3 in. x 0.3 in. x 2 in. each) glued together to form a 35 x 30 sampling array. The sides of the plexiglass blocks were coated with aluminum paint to increase the side reflectivity of the channels, and the front and back surfaces of the assembled array were polished. Ground glass was placed over the front (or sampling function side) to act as an image diffuser and over the back (or printing function side) to act as an image plane. Various printing and sampling functions were produced by superimposing the desired aperture functions over the appropriate side of the display. Simple apertures were made by machining thin aluminum sheet stock to the required geometry; more complicated apertures can be made on glass plates using photographic techniques.

The processor is operated by projecting an unsampled (analog) image onto the sampling function side of the display. The resultant image produced on the printing side of the display can be either viewed directly or photographed. The technique has the virtue of being able to produce sampled images in a large format, at high brightness, and in real time.

Figure 30 shows an original image that was used to produce the processed pictures shown in Figs. 31 and 32. Figure 31 is an example of a picture produced with full-width (s = 1) block sampling and full-width (p = 1) block printing. Figure 32 is the same image with the same number of elements, but with full-width (s = 1) block sampling and circular-aperture printing with

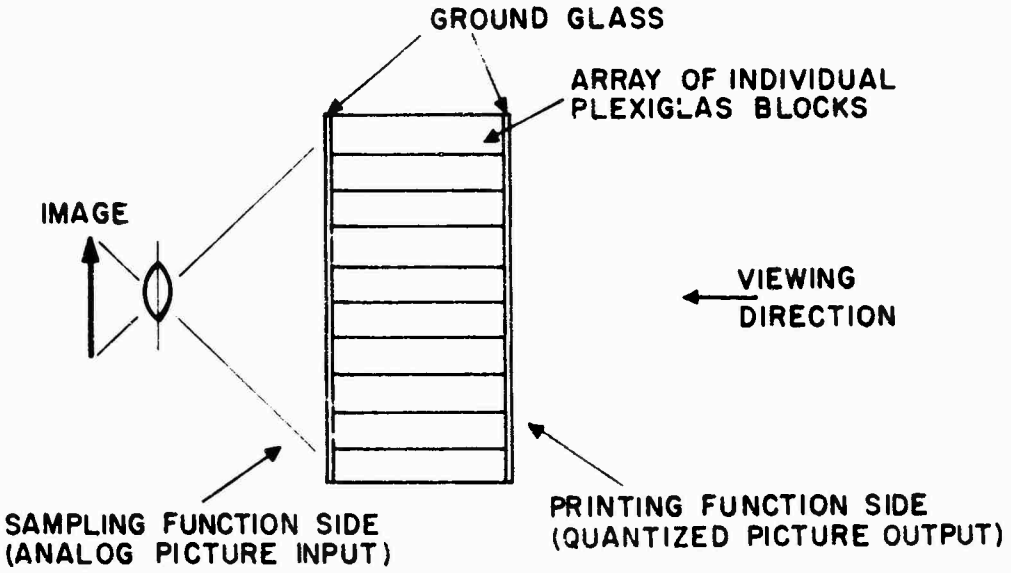


Figure 28. Side view of the optical block processor.

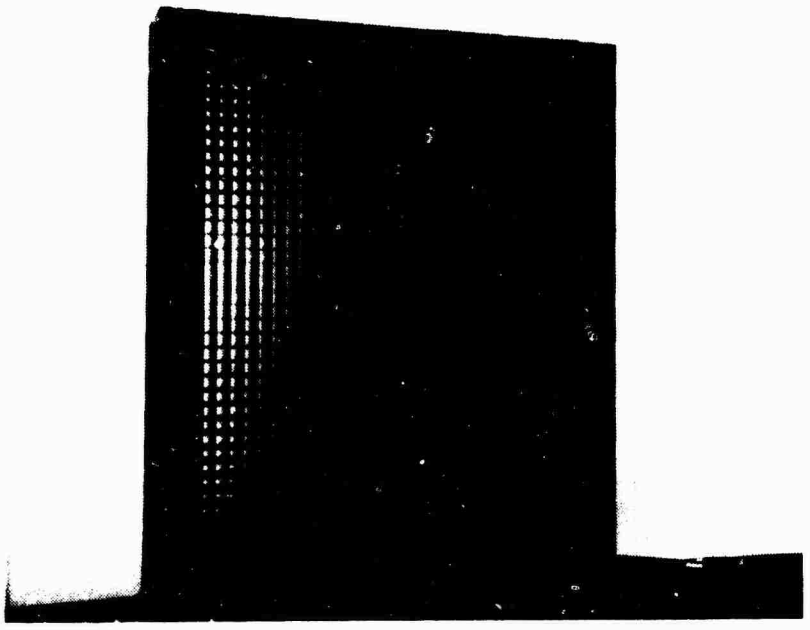


Figure 29. Photograph of the 1050-element analog block processor. The sampling and printing apertures and the ground glass image planes have been removed for the photograph.

Figure 30. Photograph of the analog input used to produce the optically processed images shown in Figs. 31 and 32.

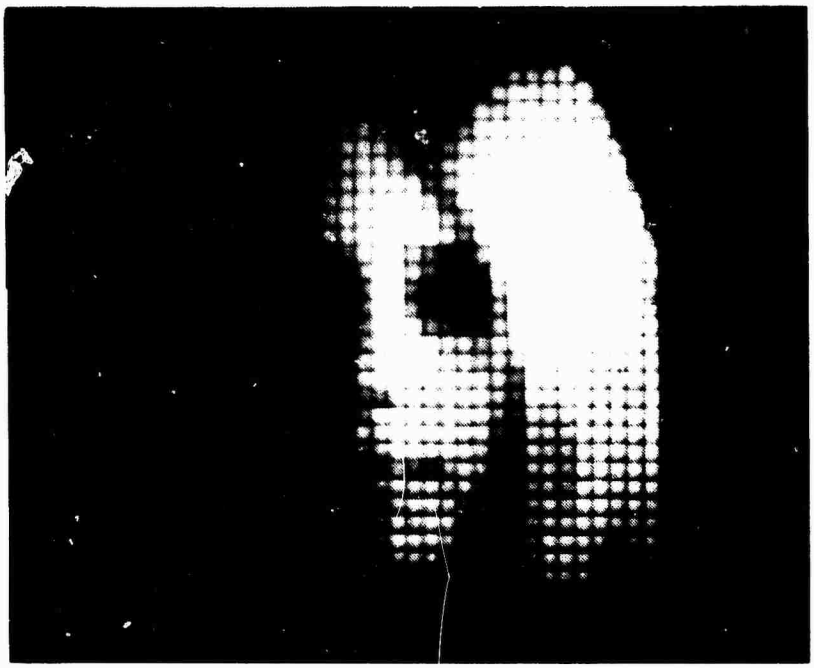


Figure 31. Sampled image produced from the image shown in Fig. 30 using the analog optical block processor with full-width block sampling and full-width block printing.

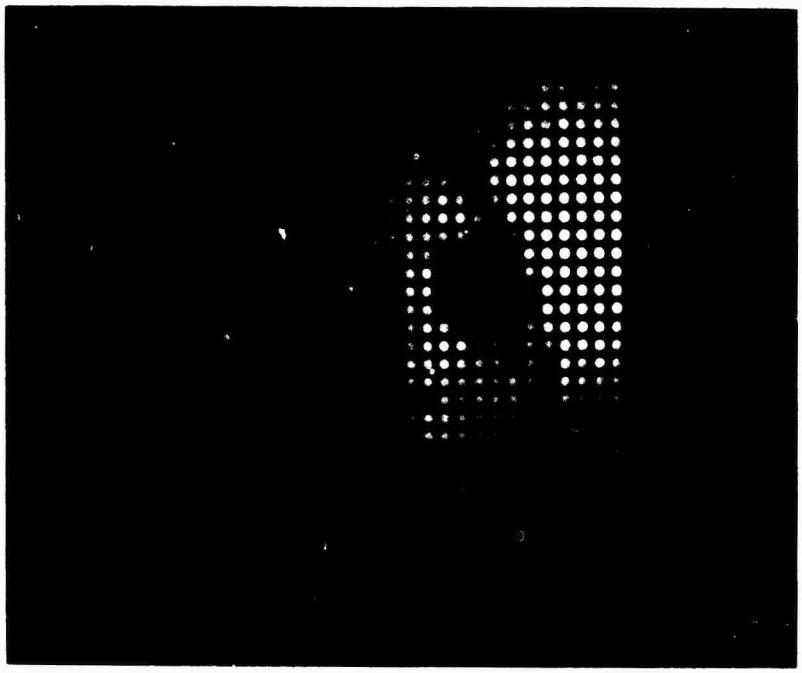


Figure 32. Sampled image produced from the image shown in Fig. 30 using the analog optical block processor with full-width block sampling and circular-aperture printing.

a ratio of active area to total block area of approximately 0.5 ($p \tilde{=} 0.7$). Although these pictures have been badly distorted by the photographic process used to produce this report, the general properties of C_v^T, S/N, H, and r_{opt} can be clearly demonstrated. Most observers judged Fig. 32 sharper but noisier than Fig. 31. Most observers also stated that they could obtain more information from pictures similar to Fig. 31 than from those similar to Fig. 32, and most found that the optimum viewing distance for Fig. 31 is somewhat smaller than that for Fig. 32. These results are in qualitative agreement with the analytic predictions of Sections III.D and E.

Note that the conventional approach to sampling noise in images suggests that the optimum viewing distance occurs where the eye acuity limit coincides with the sample spacing. In Figs. 31 and 32, the sample spacings are identical; only the printing functions are different. Yet viewer response to the two pictures in terms of perceived sharpness and noise was entirely different and in accord with the general observations arising out of our analysis.

Sampled pictures covering a wide range of sampling and printing functions can be produced to test the general validity of the image descriptors described in this report. Whereas our preliminary experimental results are in good agreement with the results of Section III, further tests are required to quantify the experimental observations and to include a large selection of observers responding to a statistically significant distribution of images.

SECTION VI

VISUAL PROCESSING OF COMPLEX TWO-DIMENSIONAL GRATINGS

A. BACKGROUND

The image descriptors developed in Sections III and IV of this report depend on the applicability of linear systems analysis to the human visual system. It is known, however, that under most conditions, the human visual response is highly nonlinear. For example, the luminance response is roughly logarithmic and contains both thresholds and saturation points [23]. A useful question to ask in connection with any real system is whether, over the operating range of interest, the assumption of linearity is valid. Some experimental success along these lines has been achieved by Blakemore and Campbell [24] and by Campbell and Robson [25] with square wave gratings. However, it has been reported that the application of linear analysis to visual problems seems to fail at low spatial frequencies [26]. (This conclusion is inconsistent with the results of the experiments presented here.)

Since the first measurement of the human visual MTF's by Schade [27], a number of other observers have performed similar experiments using simple gratings at different orientations (i.e., with the gratings running vertically, horizontally, at 45°, etc). Surprisingly, the measurement of complex two-dimensional gratings (to be defined) has not been attempted, nor have the ramifications of these measurements for two-dimensional linear processing been evaluated. In order to have a degree of confidence that two-dimensional image quality descriptors, such as that proposed in Section IV, can be applied successfully, it is necessary to determine the applicability of the concept of two-dimensional linear processing by the human visual system. In this section we summarize the results of experiments that were aimed at obtaining information along these lines.

The visual MTF is the nexus between the human observer and the application of linear systems analysis to visual problems. Because of its importance, several caveats should be mentioned about its properties. First, the spatial frequency response of an observer can be conveniently obtained by determining the value of contrast necessary for visibility at threshold for a series of

sine wave gratings at different spatial frequencies. If the threshold contrast sensitivities are inverted and plotted as in Fig. 33, a graph results which is very similar to an optical MTF. The analogy is not rigorously correct, however, because in the case of a lens one would maintain a constant input while monitoring the resultant output. When the MTF of the eye is measured using threshold techniques, just the opposite is attempted. The input contrast is varied, and it is assumed that the system is linear and that the output is constant. It should be pointed out, however, that the results of threshold measurements are similar in form to those of brightness-matching and contrast-matching experiments [5] performed well above threshold. Second, among other factors, the visual MTF is a function of the average luminance and the temporal properties of the image. The results given in this report are for a single value of average brightness and for a single set of temporal conditions. Therefore, in the future, it is obligatory that these studies be extended throughout the space-time-brightness matrix of the visual MTF.

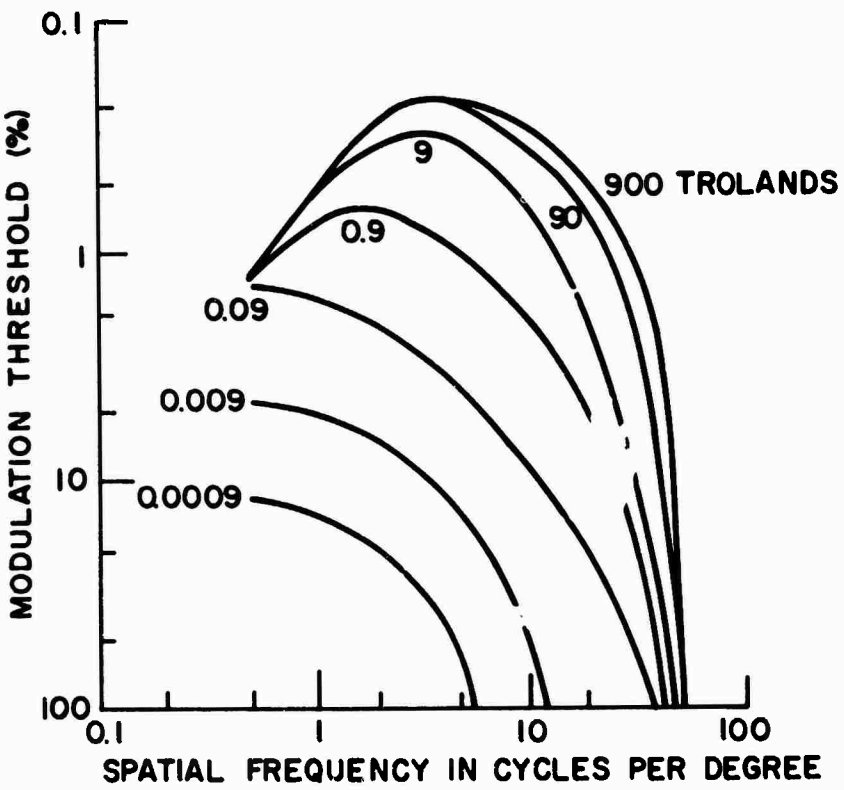


Figure 33. Modulation sensitivity threshold of the human eye in monochromatic light (λ = 5250 Å) with a 2-mm pupil as a function of retinal illumination (from ref. [28]).

B. STATEMENT OF THE PROBLEM

The experiment reported here consisted of two parts. First, the threshold sensitivity of the observer was measured as a function of retinal frequency ν for one-dimensional spatial frequency gratings oriented at various angles ϕ with respect to the horizontal as shown in Fig. 34. That is, for the one-dimentional brightness stimulus,

$$I(x) = L_o[1 + b \cos (2\pi\nu x'/r)] \tag{137}$$

it is possible to measure the value of b necessary for threshold perception at each value of ν and ϕ. From this measured b one can calculate the contrast sensitivity as shown in Fig. 35 by defining the contrast at any modulation level as

$$\text{CONTRAST} \equiv \frac{L_{max} - L_{min}}{L_{max} + L_{min}} \tag{138}$$

We obtain the visual modulation transfer function $O(\nu,\phi)$ by inverting the threshold contrast given in Eq. (138). In practice, $O(\nu,\phi)$ can be determined by making relatively few measurements at different ϕ because, as we shall see, $O(\nu,\phi)$ is nearly isotropic. The second part of the experiment required the measurement of the contrast sensitivity functions for the complex gratings

$$I(x,y) = L_o [1 + b \cos (2\pi\nu_x x/r) \cos (2\pi\nu_y y/r)] \tag{139a}$$

$$= L_o \left[1 + \frac{1}{2} b\{\cos[2\pi(\nu_x x+\nu_y y)/r] + \cos [2\pi(\nu_x x-\nu_y y)/r]\}\right] \tag{139b}$$

This intensity pattern simultaneously contains terms in all four quadrants of the two-dimensional spatial frequency space. After applying the contrast sensitivity function, obtained in the first part of our measurement program, to the Fourier transform of I(x,y) and transforming the results back to real space, it can be shown that, if $O(\nu,\phi) = O(\nu,\pi-\phi)$, the perceived intensity component proportional to b is equal to

$$E(x,y) = L_o b\, O(\nu_{eff},\phi_o)\cos(2\pi\nu_x x/r)\cos(2\pi\nu_y y/r) \tag{140}$$

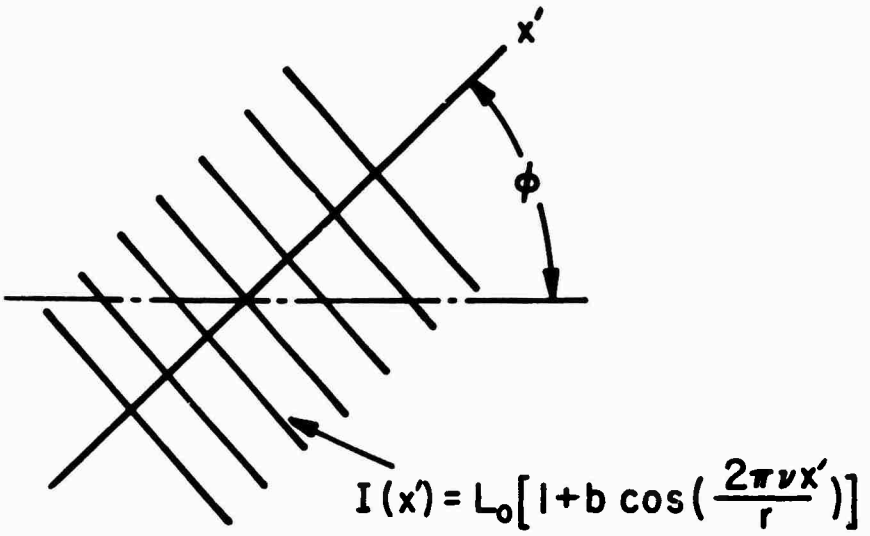
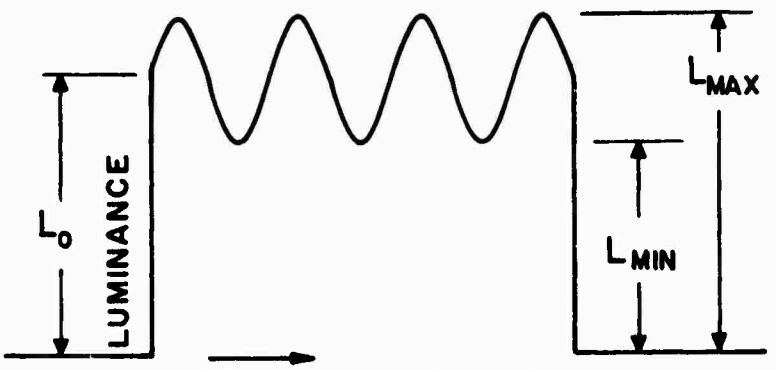


Figure 34. Definition of the polar coordinate ϕ for simple gratings oriented along the x'-axis.



$$\text{CONTRAST} = \frac{L_{MAX} - L_{MIN}}{L_{MAX} + L_{MIN}}$$

Figure 35. Definition of the contrast function.

where

$$\nu_{eff} = \left[\nu_x^2 + \nu_y^2\right]^{1/2} \qquad (141a)$$

and

$$\phi_o = \tan^{-1}\left(\nu_y/\nu_x\right) \tag{141b}$$

Thus, the expected contrast sensitivity function measured for the stimulus described in Eq. (139a) is exactly the same as for a simple one-dimensional grating of frequency ν_{eff} oriented at an angle ϕ_o from the horizontal. For the cases experimentally reported here, $\nu_x = \nu_y$, so that Eq. (140) predicts that the contrast sensitivity function measured with the complex grating would be shifted toward the low spatial frequencies by a factor of $\sqrt{2}$, when compared with one-dimensional gratings at 45°.

Figures 36 and 37 are photographs of horizontal and vertical cosine gratings which were multiplicatively combined to produce the complex grating, given in Eq. (139a) and shown in Fig. 38. Although these photographs were taken at high contrast, the general properties of the frequency shift and the angular rotation, as outlined above, are visible.

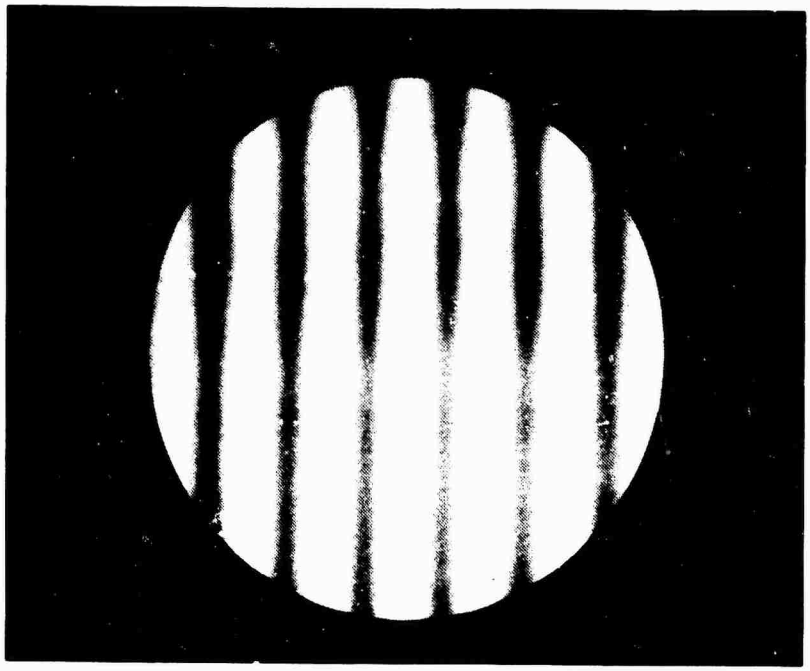


Figure 36. Horizontal cosine wave grating, $I(x) = L_o[1 + b \cos(2\pi\nu_x x/r)]$.

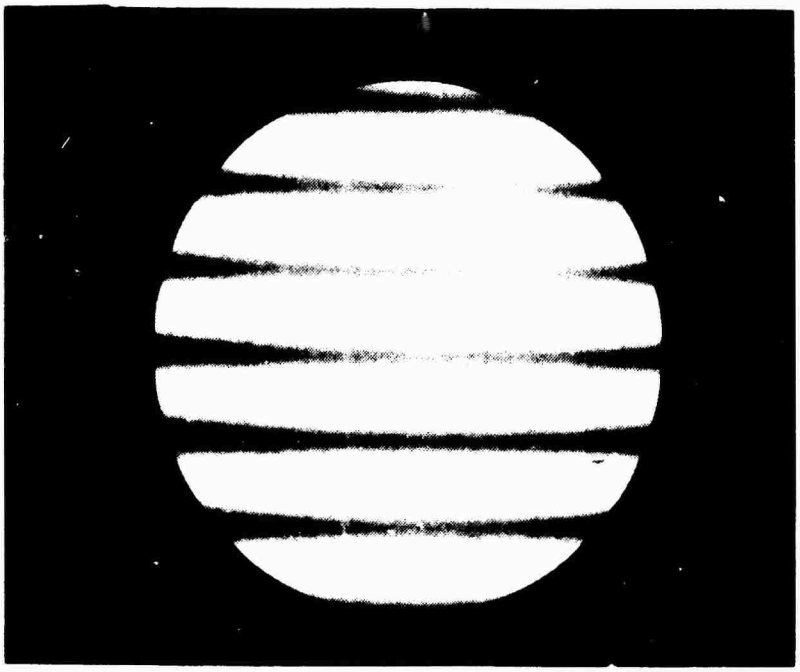


Figure 37. Vertical cosine wave grating, $I(y) = L_o[1 + b \cos (2\pi\nu_y y/r)]$.

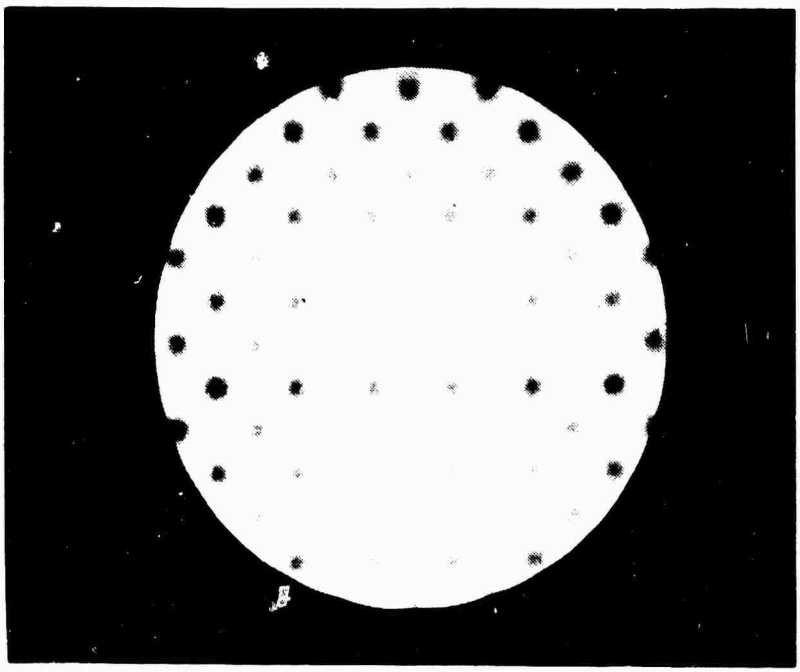


Figure 38. Complex cosine wave grating, $I(x,y) = L_o[1 + b \cos (2\pi\nu_x x/r) \cos (2\pi\nu_y y/r)]$.

EXPERIMENTAL APPROACH

The contrast sensitivity functions $O(\nu,\phi)$ were measured by determining the threshold contrast required for pattern perception. The difficulty in utilizing the results of threshold experiments arises in extrapolating the results obtained to suprathreshold levels. This remains an unsettled question in psychophysics, although it is known that under certain conditions the threshold effects are quantitatively different from the suprathreshold results [29]. It is possible to obtain visual MTF's by other techniques [5], but these techniques were not easily applicable to these studies. However, our measurements were self-consistent; that is, the visual MTF's obtained by threshold techniques were used to predict the threshold response to complex gratings.

A set of preliminary experiments were performed to determine some of the general properties of the visual MTF and to determine the level of experimental sophistication necessary to perform the experiments with sufficient accuracy to resolve the phenomena of interest. Although the details of these experiments were given elsewhere [30], we list here the following conclusions:

(1) Any attempt to measure a single visual MTF with adequate precision requires a large number of data points (typically greater than 10^3 binary decisions). Therefore, measurements of the contrast sensitivity function of several observers at the required values of ϕ with both simple and complex gratings requires approximately 10^5 data points. If the experiments are performed manually, this requires data acquisition times and data manipulations that are prohibitive. The situation becomes even more prohibitive when parametric studies are performed to determine both the background brightness dependence and the temporal properties of the visual MTF.

(2) The most suitable experimental approach is a forced-choice procedure [31]. This procedure is readily automated, and it has been found to yield highly reproducible results.

(3) The only practical approach to these experiments is to develop an all-electronic system with interactive computer control.

D. EXPERIMENTAL STUDIES DESIGN

Our objective was to develop an experimental facility that was flexible and totally automated; i.e., to develop a system that would allow the data to be acquired quickly and without the need for manual data manipulation. The block diagram of the system developed is shown in Fig. 39.

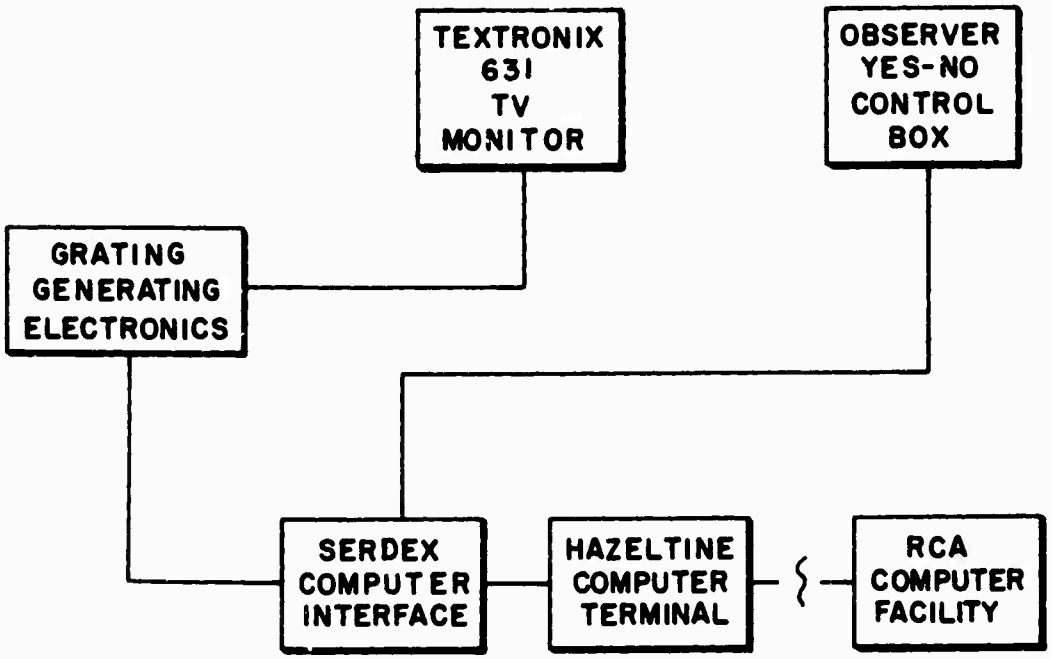


Figure 39. Block diagram of the experimental apparatus.

A high-quality Tektronix 631 monitor was utilized as the display device. Even high-quality monitors do not represent a perfect display device; they have a raster line structure which is slightly visible at typical viewing distances, they exhibit some 30-cycle flicker at high brightness levels, their output is slightly noisy, the uniformity of their edge-to-edge brightness is not perfect (see Fig. 40), and the total effective aperture size is only marginally sufficient (typically around 7° of vision for experiments of this type). The advantages of using a monitor are that (1) it allows the generation of a wide variety of gratings; (2) the display is bright; (3) it is easily calibrated; (4) it has a large dynamic range; and (5) it is an all-electronic system that allows easy automation.

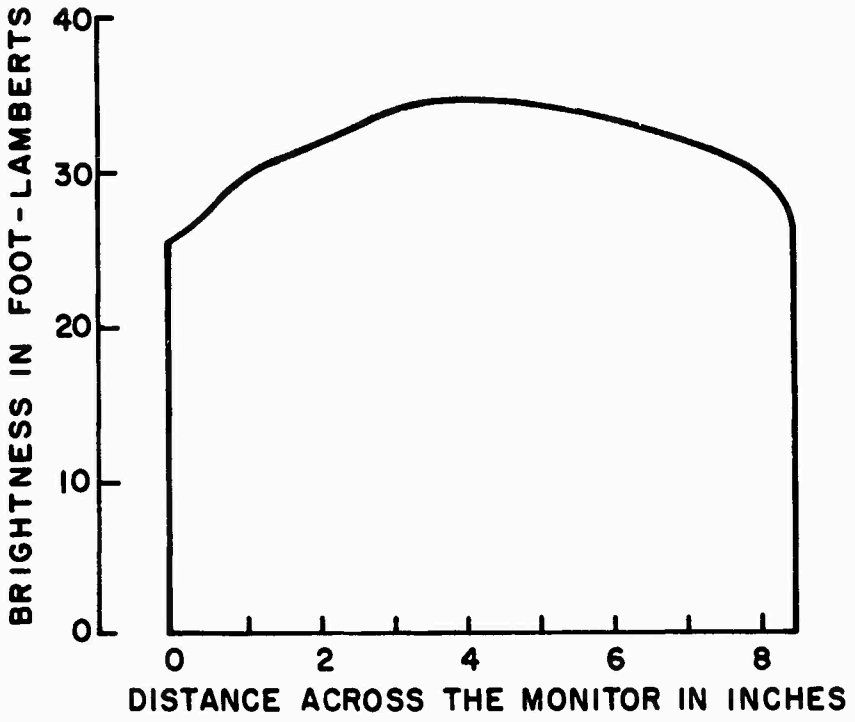
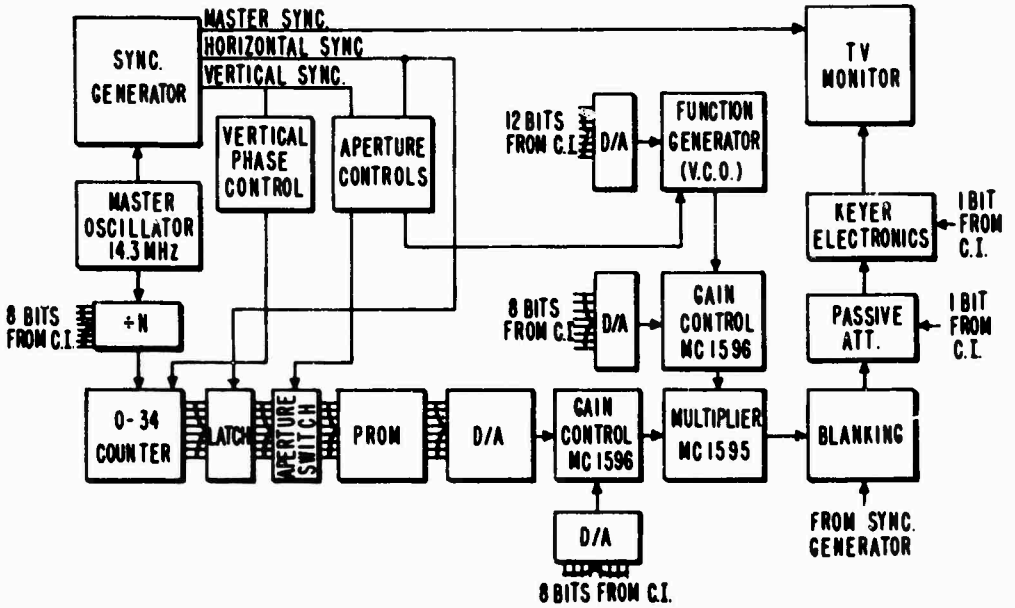


Figure 40. Edge-to-edge brightness profile of the monitor along the horizontal direction (along the raster lines) in the center of the vertical plane.



C. I. = computer interface.

Figure 41. Detailed block diagram of the two-dimensional sine wave generating apparatus.

A detailed block diagram of the system is shown in Fig. 41. Because the display produces a sampled output in the vertical direction, it was necessary to generate the vertical gratings in a manner that minimized sampling errors and Moiré effects. Sampling errors become a particularly serious problem for high-frequency gratings because the number of samples per cycle is small. A convenient way to generate stable gratings with the proper phase relationships was to count down from the master clock oscillator of the sync generator which is used to produce the horizontal and vertical sync pulses to operate the monitor. The master oscillator output is first divided, then the divided output drives a counter. The counter produces a 7-bit binary output which cycles between 0 and 34. This output is used to drive a 35-word by 8-bit programmable read-only memory (PROM). This element simply converts the input count into the proper binary output for a sine wave. The output from the PROM is converted to analog form by a suitable D/A. By dividing down properly from the master oscillator, any vertical frequency up to the sampling limits of the monitor can be produced. We found that it is possible to produce a very acceptable sine wave by using only two samples per cycle, that is, by having one horizontal line on, and one horizontal line off. This is possible because the beam profile on the monitor is approximately Gaussian, and the higher harmonics which are missing in such a poorly sampled grating fall outside the range of visual perception. In practice, however, this frequency cannot be utilized because of the interlace used to produce the picture on a monitor. Since every other line is scanned per half frame, this means that at the highest spatial frequency (which is produced by having successive bright and dark horizontal lines), the grating is formed by scanning all the high brightness lines prior to the darker lines. This results in an objectionable amount of 30-cycle flicker. We found that at least five samples per cycle were necessary to vitiate the observable flicker in the vertical direction. Since the digitally derived sine wave was only 35 words long, very low spatial frequencies exhibited some contouring at high contrast settings. However, at threshold the contouring was never visible (see, for example, Fig. 37). Following the 0 to 34 counter, a latch was used to maintain the sampled value of a given vertical sine wave over one complete horizontal scan line. The output from the latch was then fed to an aperture switch which was used to control the number of cycles displayed on the monitor at a given spatial frequency. (In these studies the maximum aperture of the monitor was used at all times.)

In the horizontal direction, the monitor is strictly an analog device, and it was not required to digitally derive the gratings. The voltage-controlled oscillator input of a Wavetek Model 162 function generator was accessed through a 12-bit D/A from the computer interface to produce sine waves from 50 kHz to 5 MHz. Each horizontal line was triggered using the horizontal sync pulses from the sync generator to obtain the proper phasing.

Control over the output amplitude was provided by using Motorola MC1596 (horizontally running gratings) and MC1595 (vertically running gratings) multipliers as linear gain controls. These devices were controlled through 8-bit D/A converters from the computer interface. Additional passive attenuators were used following the linear gain controls to obtain the required dynamic range while maintaining a minimum peak-to-peak brightness modulation signal-to-rms-electronic-noise-ratio of 20 at the lowest contrast settings. When generating complex gratings, an additional MC1596 was used to perform the multiplication $\cos(2\pi\nu_x x/r)$ times $\cos(2\pi\nu_y y/r)$. Finally, the signal was amplified and buffered, the appropriate blanking added, and then sent through a keyer to the monitor. The keyer was controlled by the computer and allowed individual gratings to be displayed for a predetermined length of time. Regardless of the state of the keyer, the average brightness on the monitor was held constant.

The computer interface was constructed from modules made by Analog Devices, Inc. with expander boards and latches for the reception of a total of 16 ASCII words (4-bits per word) and the transmission of 8 ASCII words. The computer interface was, in turn, interfaced with the RCA Laboratories Time Sharing System through a Hazeltine terminal which was located in the laboratory where the experiments were performed. Although time sharing systems do not usually allow for instantaneous response, the turn-around time was kept to less than 1 second by operating during low-demand periods.

The software was developed to run an experiment without manual control and to store the data acquired in the appropriate computer files. The computer program selected a spatial frequency, the type of grating to be displayed, and its contrast. The computer sent the appropriate ASCII characters back to the computer interface logic which, in turn, set the appropriate D/A's, relays, etc., to generate the required pattern. The pattern was displayed for a selected time interval and the observer indicated whether the grating was

above threshold or below it. He conveyed his answer by depressing a button marked "yes" or "no" on a small box which he held in his hand. Two aural indicators were used to orient the observer to the presence of a grating. One-half of a second before a grating was flashed on the display, a brief "bleep" was sounded to indicate that a grating was about to be presented. Then, while the grating was on the screen, a faint 1-kHz tone was sounded to indicate the presence of the grating. After the tone stopped, the observer indicated whether he saw the grating or not, and his answer was immediately transmitted back to the computer. (Without the aural indicators, the observer was easily confused and frustrated by the experimental situation because he did not know when a grating would appear.) A complete sequence for each grating, including the computer selection display presentation and observer response, took approximately 3 seconds. After receiving the observer's answer, the computer program selected another frequency, mode, and contrast setting, and the procedure was repeated. The frequencies and modes were selected at random using a card shuffling routine, and the intensities were selected using a simple "up-down" algorithm that was controlled by the observer's previous answers.

The monitor was mounted as shown in Fig. 42. A circular surround was mounted over the monitor which gave a total angular field for the display of 6.5° at a viewing distance of 190 cm. The area forward of the observer, around the display, was covered and painted with flat white paint. The experimental area was lightly illuminated, with the brightness of the surround measuring 3.5 ft-L. Partitions were placed on either side of the observer so that his visual field could be controlled. The actual experimental environment is shown in Fig. 43 with an observer holding the "yes-no" box.

The monitor-electronics display system presented three different types (or modes) of gratings at a single value of ϕ, as shown in Figs. 36, 37, and 38; that is, in this case, gratings running along the horizontal, 90° to the horizontal, and the product of these two. For other angular orientations, the monitor was physically rotated. A single experiment consisted of determining all of the contrast sensitivity threshold values for all three modes and at all frequencies for a given rotation of the monitor.

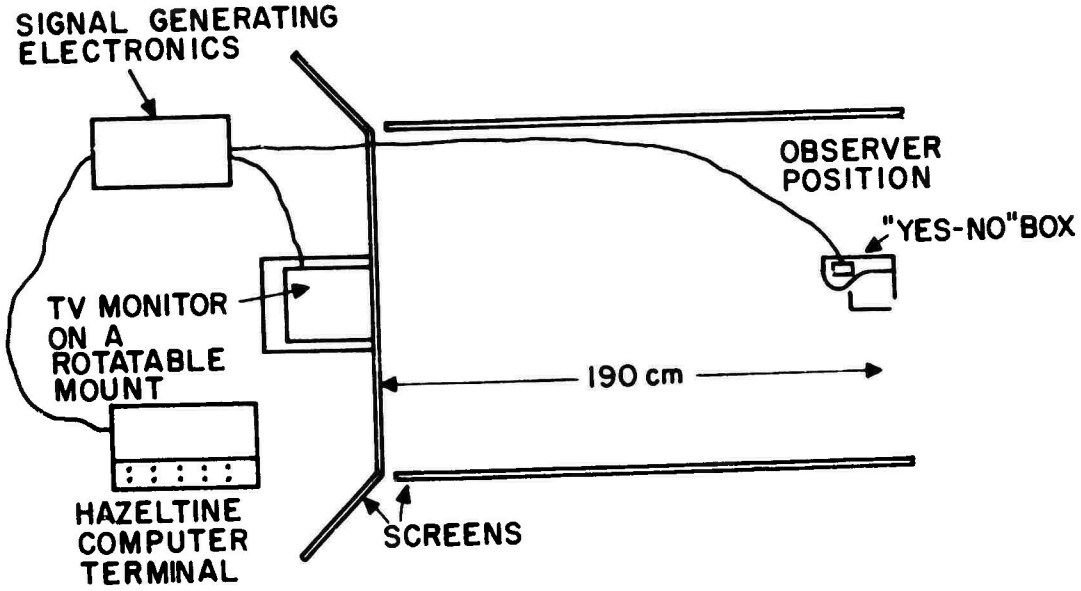


Figure 42. Schematic diagram of the experimental situation. The viewing distance is defined as the distance between the monitor screen and the observer's eyes.

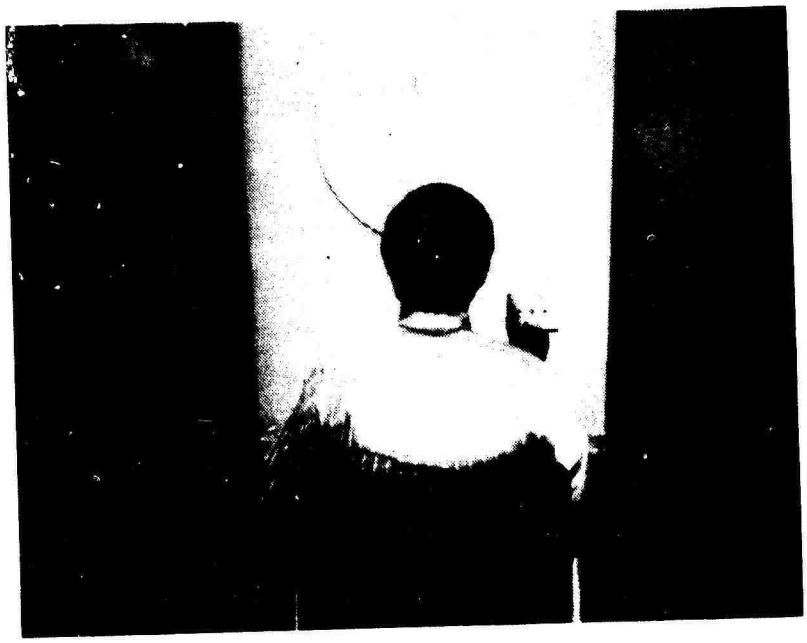


Figure 43. Experimental environment showing an observer holding the "yes-no" box.

E. CALIBRATION

The monitor and associated electronics were calibrated using a Spectra Pritchard Photometer Model 1980. A special rectangular aperture was used with the photometer that allowed the highest spatial frequencies displayed on the monitor to be resolved while averaging over several vertical raster lines. The monitor was calibrated both vertically and horizontally using the photopic response curve of the photometer with a reproducibility of better than 5%. It was found that the monitor's output was linear with respect to the input voltage for contrasts up to 50% at an average luminance of 33 ft-L. No measurements were made at contrasts above this contrast setting.

F. EXPERIMENTAL PROCEDURE

For each rotation of the monitor, three modes [simple gratings at a given ϕ, simple gratings at $(\phi + 90°)$, and complex gratings composed from these two simple gratings] and 31 different spatial frequencies were displayed during a typical experiment. The spatial frequencies chosen were distributed logarithmically and covered a range of approximately 100 to 1 at a given observer location. One hundred individual intensity steps were provided at 7% intervals which covered contrasts from 0.06% to 50%. [The results of these measurements indicated that smaller intervals ($\sim$3 to 4%) would be more appropriate.] Four rotations of the monitor were utilized at 0°, 30°, 60°, and 90° from the horizontal. The display was circular with an average luminance of 33 ft-L. The surround was held at a constant 3. ft-L, and the display time for a grating was 1.0 second. The display time was alternated with approximately 2 seconds of blank field at the same average luminance as that during the grating presentation. Because the entire set of data took many hours to collect, runs were broken into 15-minute sessions with rest intervals between sessions.

Artificial pupils and monocular vision were not used in these experiments because an attempt was made to replicate normal viewing conditions. All the experimental runs were begun below threshold at each frequency and mode and progressively moved into the threshold regime. A fixation point was not used.

A single observer (R.I.) was tested in these experiments. He was 38 years old and was without any known vision abnormalities. He had recently consulted

an ophthalmologist and had a new prescription to correct for a slight case of near-sightedness. He was well trained in the use of the apparatus and had been tested for about 10 hours before the results reported here were obtained.

G. RESULTS AND DISCUSSION

The contrast sensitivity functions for simple one-dimensional sine wave gratings are shown in Figs. 44 through 49. This information is replotted in Figs. 50, 51, and 52 to show more clearly the effects of grating orientation on the contrast sensitivity functions. All of these curves exhibit the characteristic shape of the visual contrast sensitivity function [28]. At low spatial frequencies the threshold required for perception rolls off approximately linearly with decreasing spatial frequency; at high spatial frequencies it rolls off approximately as the square of the spatial frequency. The spatial frequency for maximum contrast sensitivity in all cases is between 3 and 4 cycles/degree-of-vision, and the minimum contrast sensitivity is approximately 0.25%.

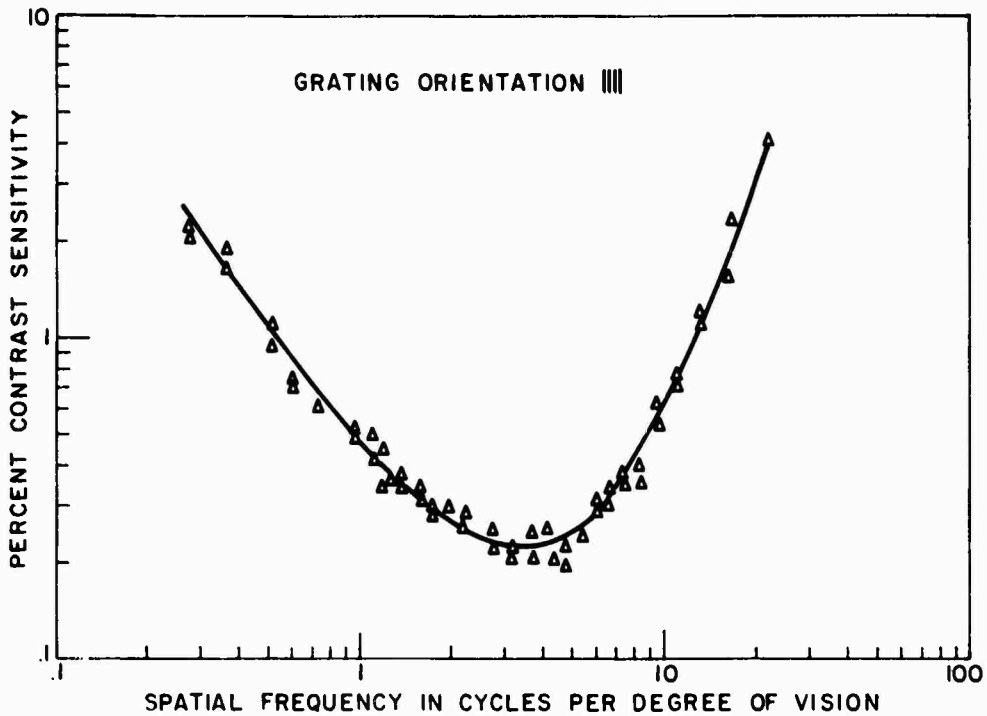


Figure 44. Threshold contrast sensitivity vs spatial frequency for simple gratings oriented at $\phi = 0°$.

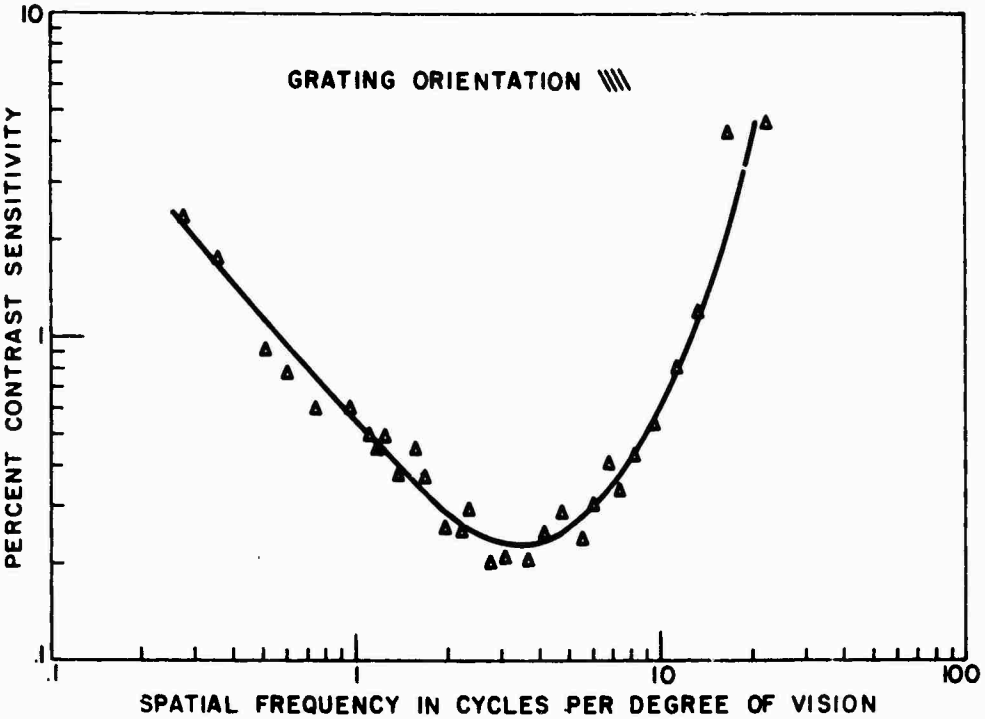


Figure 45. Threshold contrast sensitivity vs spatial frequency for simple gratings oriented at $\phi = 30°$.

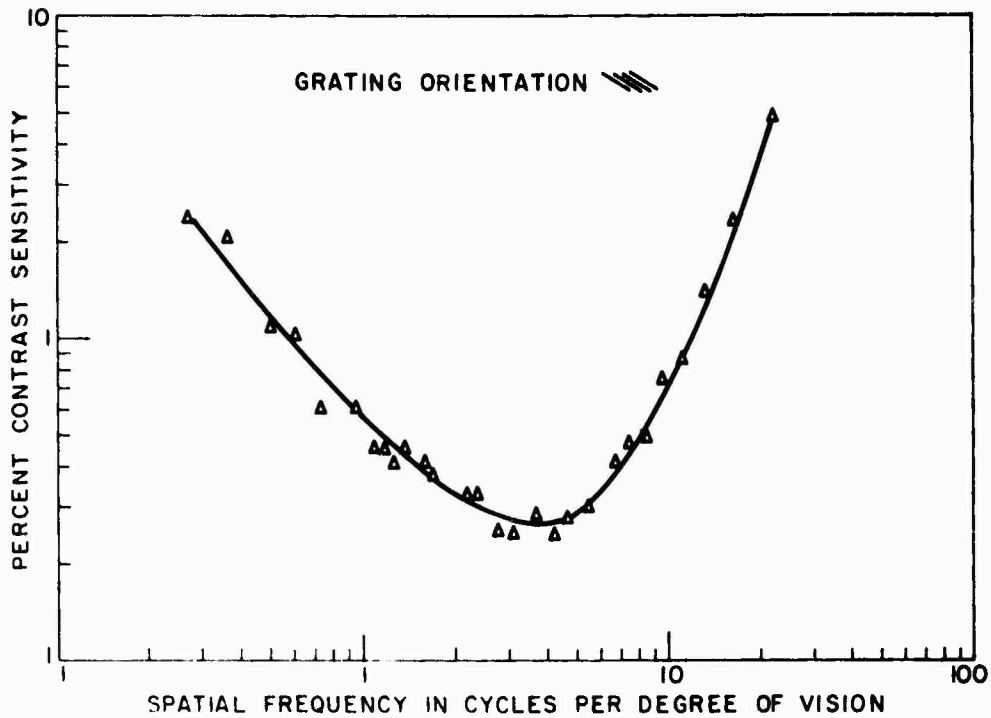


Figure 46. Threshold contrast sensitivity vs spatial frequency for simple gratings oriented at $\phi = 60°$.

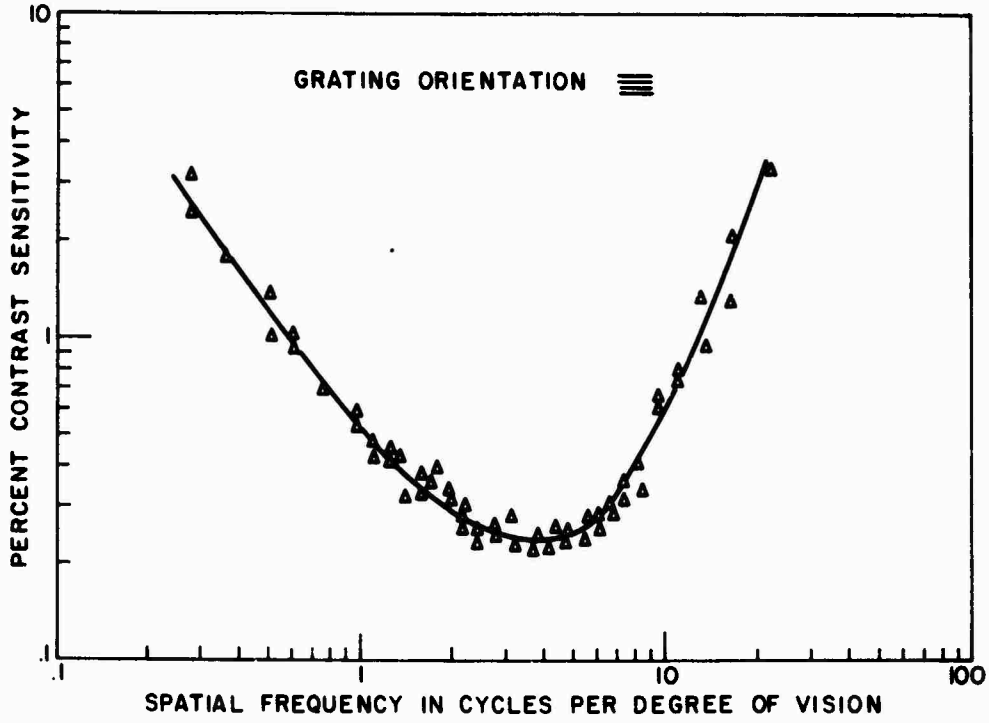


Figure 47. Threshold contrast sensitivity vs spatial frequency for simple gratings oriented at $\phi = 90°$.

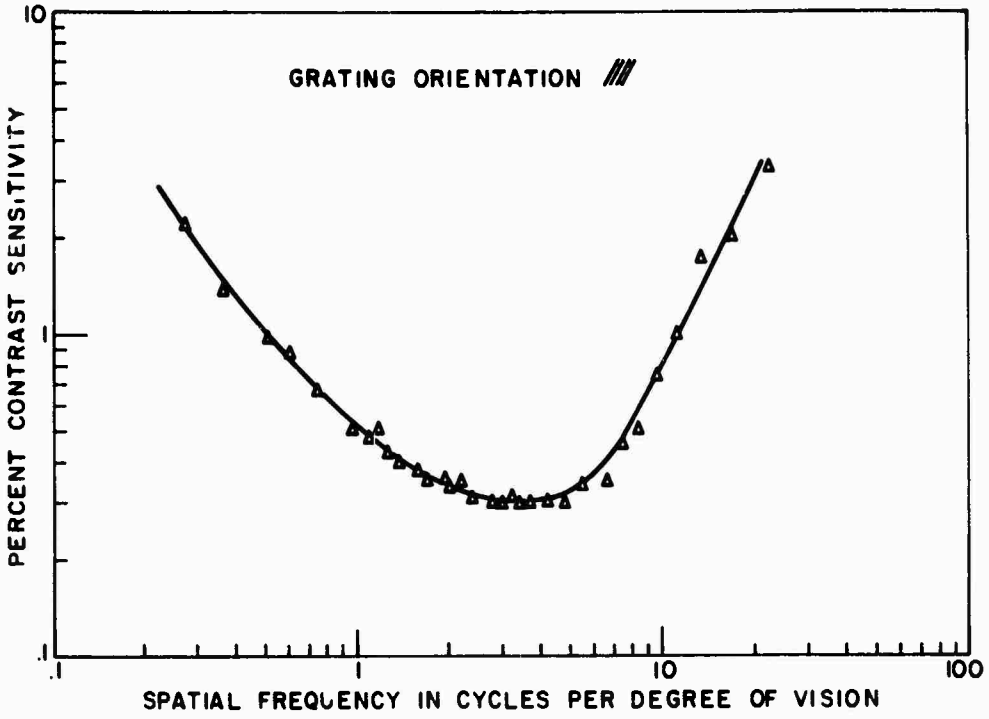


Figure 48. Threshold contrast sensitivity vs spatial frequency for simple gratings oriented at $\phi = 120°$.

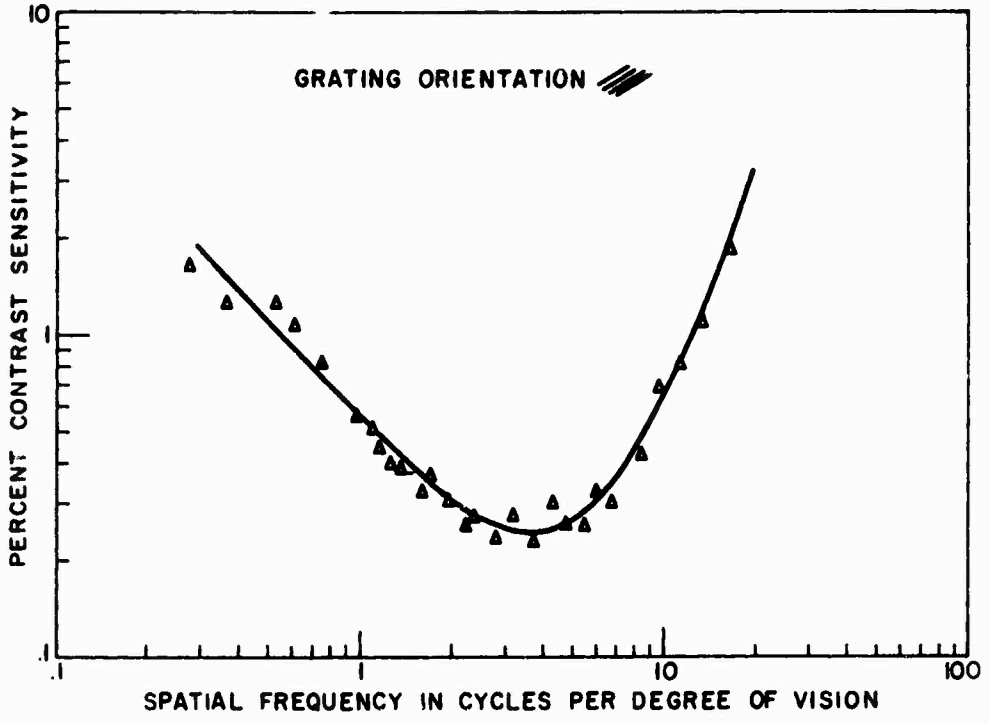


Figure 49. Threshold contrast sensitivity vs spatial frequency for simple gratings oriented at $\phi = 150°$.

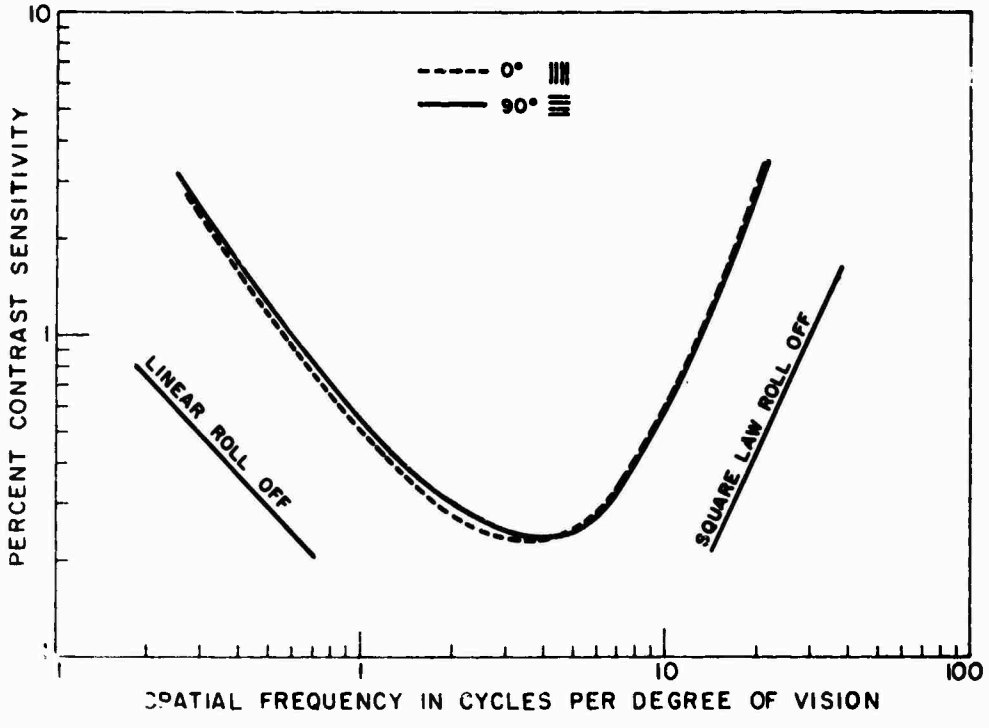


Figure 50. Threshold contrast sensitivity vs spatial frequency for simple gratings oriented at $\phi = 0°$ and $\phi = 90°$.

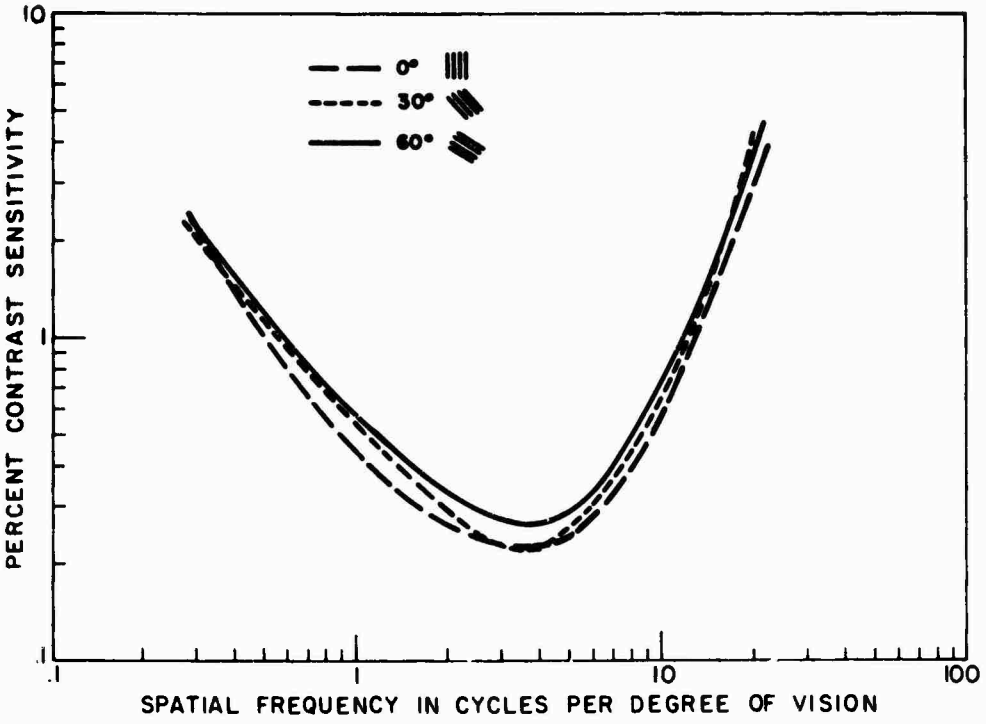


Figure 51. Threshold contrast sensitivity vs spatial frequency for simple gratings oriented at ϕ = 0°, 30°, and 60°.

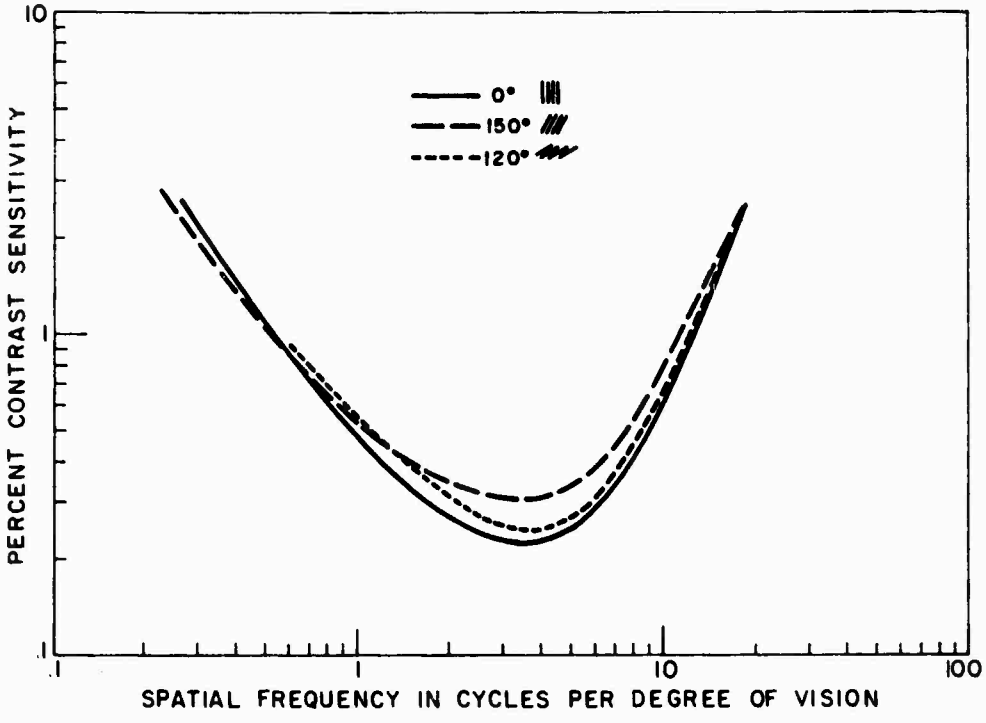


Figure 52. Threshold contrast sensitivity vs spatial frequency for simple gratings oriented at ϕ = 0°, 120°, and 150°.

Figure 50 indicates that, for this subject, the difference between his vertical and horizontal response to simple gratings is negligible. Figures 51 and 52 show a decrease in sensitivity for gratings oriented at an oblique angle to the horizontal, with minimum sensitivities at $\phi = 150°$ and $\phi = 60°$, suggesting that this subject's MTF is rotated slightly counterclockwise.

The results for the threshold sensitivity measurements using complex gratings are shown in Figs. 53, 54, and 55 (for all three cases, $\nu_x = \nu_y$). The results of these figures have been replotted in Fig. 56. These curves are actually quite similar, with the discrepancies about the same as the spread in the data. It is somewhat surprising that the gratings at $\phi_o = 45°$ should have the greatest sensitivity. On the basis of linear analysis one would predict that these gratings would have had the worst sensitivity. Perhaps more significantly, however, the peak sensitivity for all three curves is shifted below 3 cycles/degree-of-vision, whereas the peak sensitivity for the simple gratings occurred above 3 cycles/degree-of-vision.

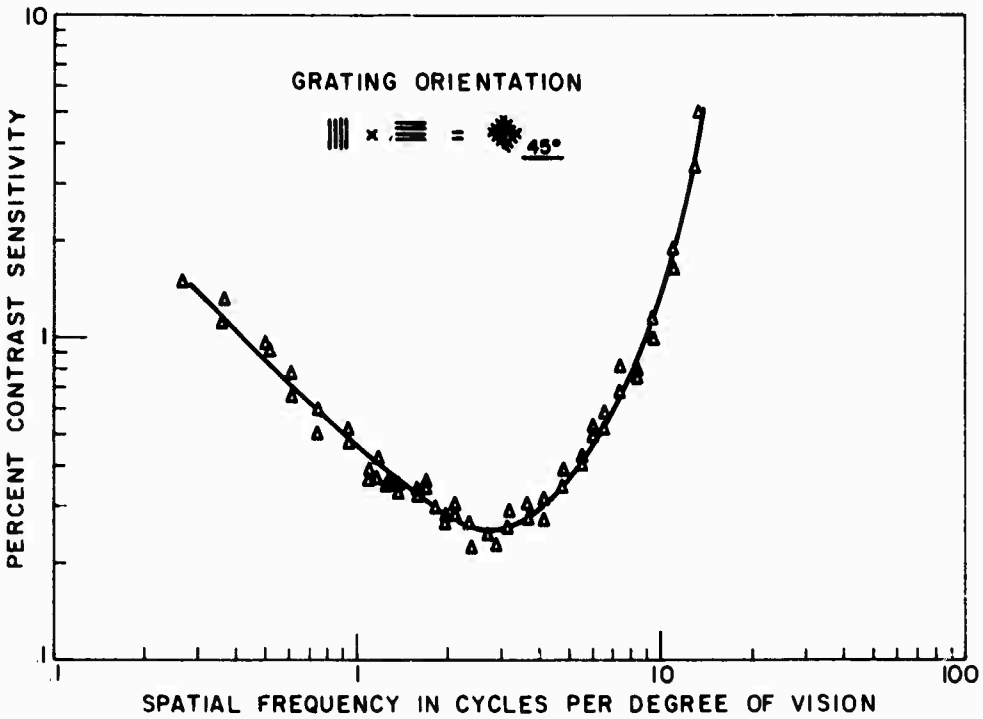


Figure 53. Threshold contrast sensitivity vs spatial frequency for complex gratings oriented at $\phi_o = 45°$.

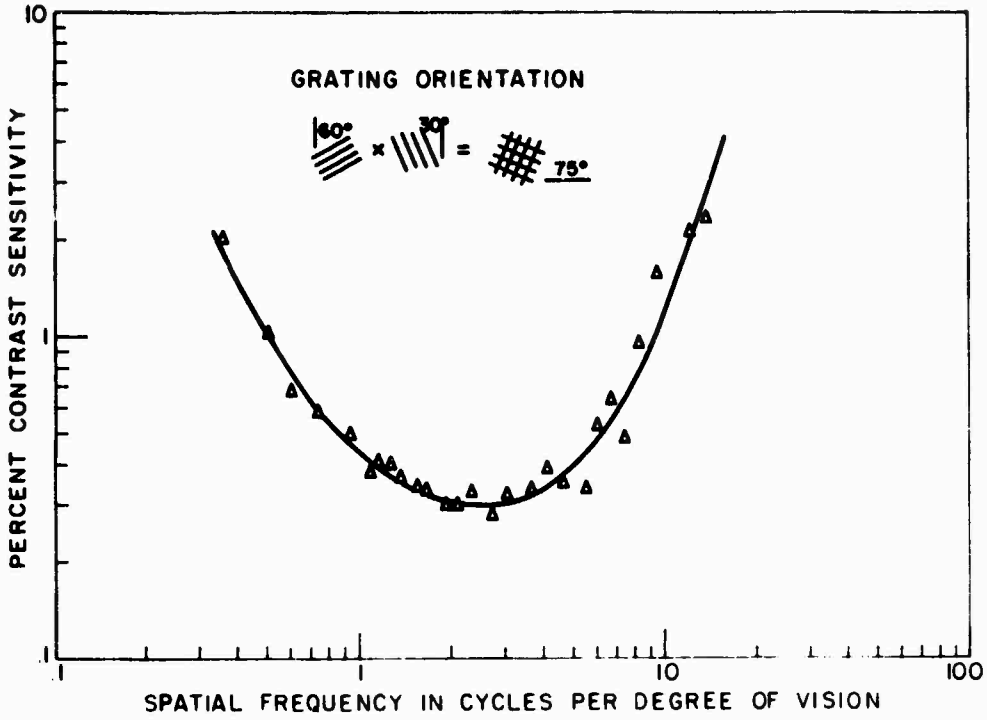


Figure 54. Threshold contrast sensitivity vs spatial frequency for complex gratings oriented at $\phi_o = 75°$.

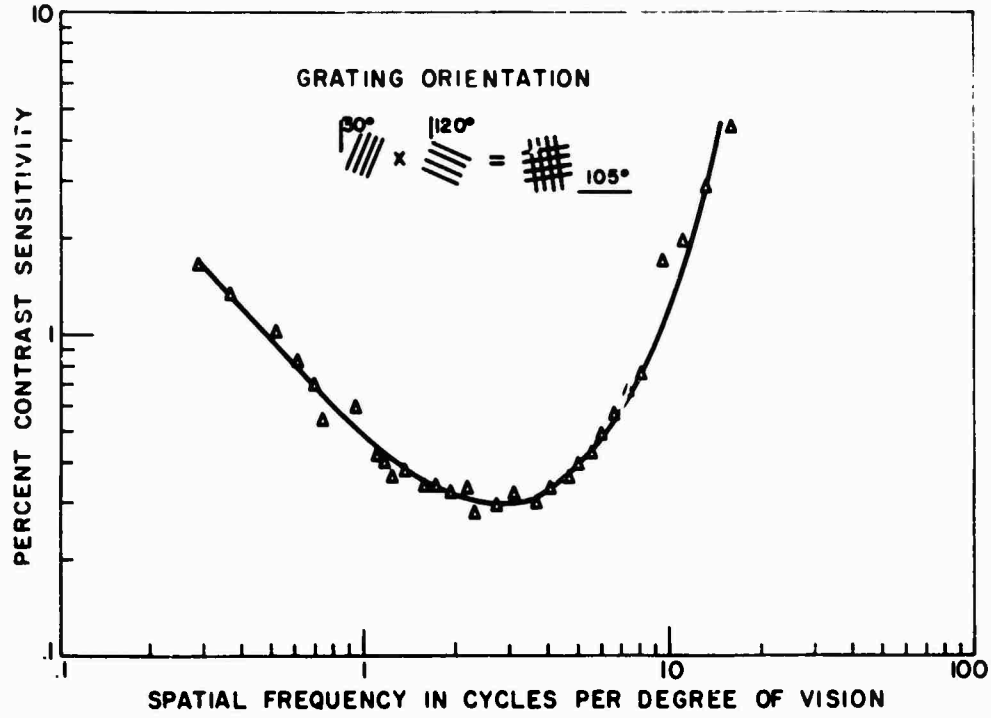


Figure 55. Threshold contrast sensitivity vs spatial frequency for complex gratings oriented at $\phi_o = 105°$.

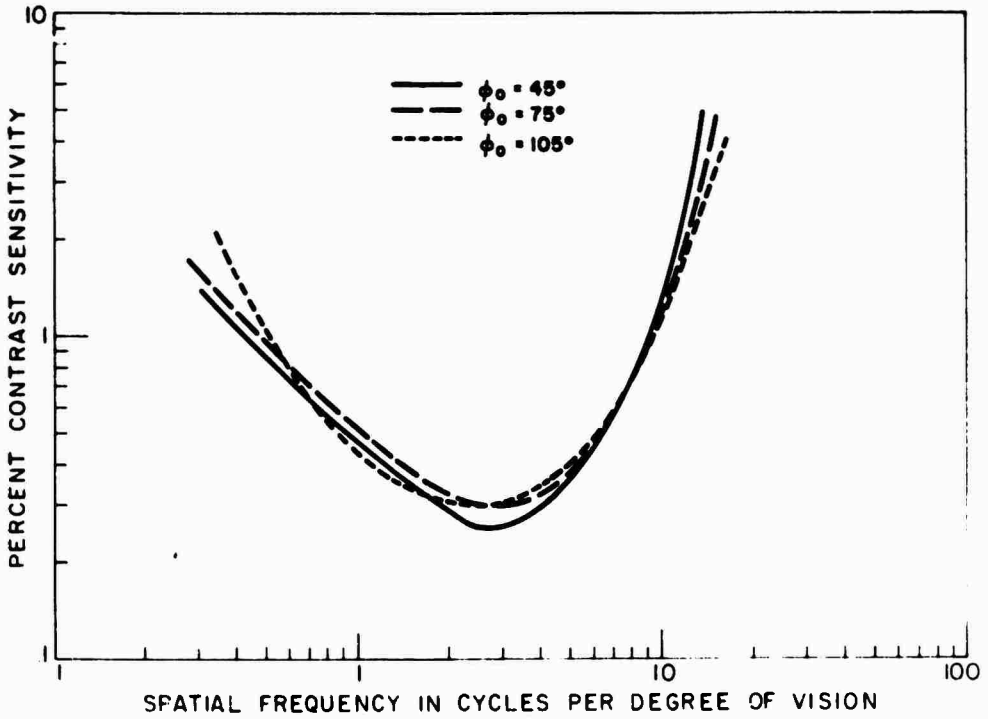


Figure 56. Threshold contrast sensitivity vs spatial frequency for complex gratings oriented at ϕ_o = 45°, 75°, and 105°.

Figures 57, 58, and 59 are a series of graphs of the contrast sensitivity versus spatial frequency for complex gratings with the predicted polar rotations* of, respectively, 45°, 75°, and 105°, and for simple gratings at the same rotations. The contrast sensitivity curves for the simple gratings at these orientations were approximated from Fig. 51. This is a rather remarkable set of curves when each set is normalized to the same peak sensitivity, because in each case the contrast sensitivity curve for the complex gratings is shifted toward the lower frequencies. This result is consistent with linear analysis. Indeed, the shift in the peak sensitivity for each of these curves is impressively close to $\sqrt{2}$, although the actual magnitude of the shift varies considerably at other spatial frequencies.

*E.g., a polar rotation of 45° implies that the complex grating was formed by multiplying together a vertical and a horizontal grating. [See Eqs. (141a) and (141b).]

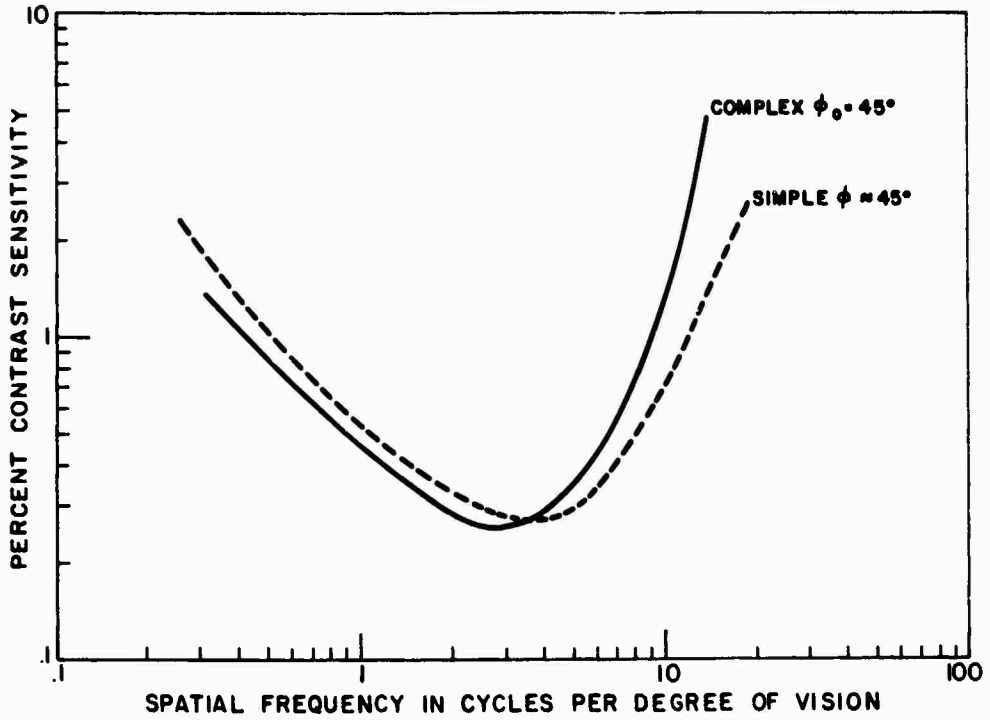


Figure 57. Threshold contrast sensitivity vs spatial frequency for simple and complex gratings oriented at ϕ = 45°.

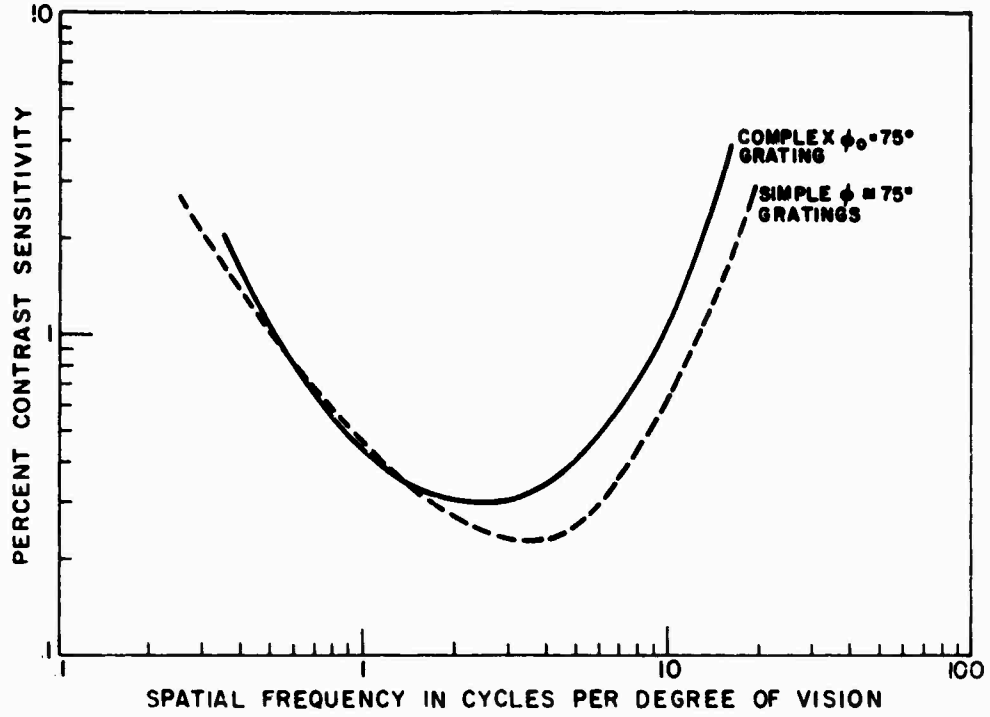


Figure 58. Threshold contrast sensitivity vs spatial frequency for simple and complex gratings oriented at ϕ = 75°.

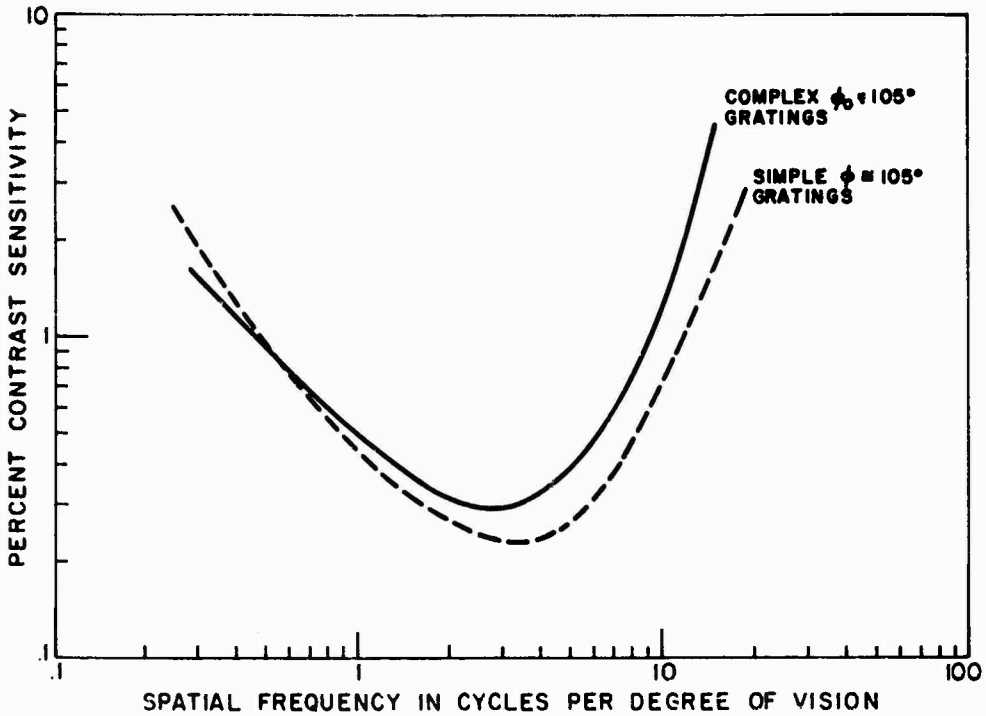


Figure 59. Threshold contast sensitivity vs spatial frequency for simple and complex gratings oriented at $\phi = 105°$.

REFERENCES

1. Mischa Schwartz, *Information Transmission, Modulation, and Noise,* (2nd Edition, McGraw-Hill Book Co., New York, 1969) Chapter 6.

2. O. H. Schade, A New System of Measuring and Specifying Image Definition," *Optical Image Evaluation,* National Bureau of Standards Circular 526, U.S. Gov. Printing Office, Washington, D. C., 1954, pp. 231-249.

3. J. B. Chatten et al., "The Composite Video Signal - Waveforms and Spectra," IRE Trans. on Broadcast and Television Receivers BTR-1, 31 (1955).

4. L. D. Harmon and B. Julesz, "Masking in Visual Recognition: Effects of Two-Dimensional Filtered Noise," Science 180, 1194 (1973).

5. The frequency response of the human visual system has been measured by methods that involve: (i) threshold direction [O. H. Schade, "Optical and Photoelectric Analog of the Eye," J. Opt. Soc. Am. 46, 721 (1956)]; (ii) brightness matching [E. M. Lowry and J. J. DePalma, "Sine Wave Response of the Visual System: 1. The Mach Phenomenon," J. Opt. Soc. Am. 51, 740, (1961); O. Bryngdahl, "Characteristics of the Visual System: Psychophysical Measurements of the Response to Spatial Sine-Wave Stimuli in the Photopic Region," J. Opt. Soc. Am. 56, 811 (1966); G. A. Hay and M. S. Chesters, "Signal-Transfer Functions in Threshold and Suprathreshold Vision," J. Opt. Soc. Am. 62, 990 (1972)]; and (iii) contrast matching [M. Davidson: "Perturbation Approach to Spatial Brightness Interaction in Human Vision," J. Opt. Soc. Am. 58, 1300 (1968)].

6. F. W. Campbell and D. G. Green, "Optical and Retinal Factors Affecting Resolution," J. Physiol. 181, 576 (1965); M. Davidson, "Perturbation Approach to Spatial Brightness Interaction in Human Vision," J. Opt. Soc. Am. 58, 1300 (1968).

7. D. H. Kelly, "Frequency Doubling in Visual Responses," J. Opt. Soc. Am. 56, 1628 (1966).

8. J. G. Robson, "Spatial and Temporal Contrast-Sensitivity Functions of the Visual System," J. Opt. Soc. Am. 56, 1141 (1966).

9. See, for example, G. J. C. van der Horst, "Fourier Analysis and Color Discrimination," J. Opt. Soc. Am. 59, 1670 (1969); E. M. Granger and J. C. Heurtley, "Visual Chromaticity-Modulation Transfer Function," J. Opt. Soc. Am. 63, 1173 (1973).

10. C. E. Shannon, "Communication in the Presence of Noise," Proc. IRE 37, 10 (1949).

11. R. W. Cohen and I. Gorog, "Visual Capacity - An Image Quality Descriptor for Display Evaluation," Proc. S.I.D. 15, 53 (1974).

12. The relation $|\hat{\Pi}(\omega)| \leq 1$ follows from Eq. (11) and the fact that $\hat{P}(\omega)$ is the Fourier transform of a positive real function $P(x)$. See R. W. Cohen and I. Gorog, J. Opt. Soc. Am. 63, 1071 (1973).

13. See, for example, E. W. Engstrom, "A Study of Television Image Characteristics," Proc. IRE 21, 1631 (1933).

14. See, for example, A. Rose, *Vision-Human and Electronic* (Plenum Press, New York, 1973), p. 33; T. N. Cornsweet, *Visual Perception*, (Academic Press, New York, 1970) pp. 84, 85.

15. H. B. Dwight, *Tables of Integrals and Other Mathematical Data* (MacMillan Co., New York, 1961) p. 12.

16. L. B. W. Jolley, *Summation of Series* (Dover Publications, Inc., New York, 1961) pp. 102, 103. Successive differentiation and rearrangement of Jolley's series (551) and (552) leads to Eq. (71).

17. The correlation quality and mean square error have been proposed by E. H. Linfoot [J. Opt. Soc. Am. 46, 740 (1956)] as quality factors for noiseless two-dimensional displays. Linfoot neglected the effect of the human visual system and assumed a white signal power spectrum. The properties of a quantity similar to ϵ, including the human visual system, have been considered independently by D. O. North (private communication).

18. H. Nyquist, "Certain Factors Affecting Telegraph Speed," Bell Syst. Tech. J. 3, 324 (1924).

19. R. V. L. Hartley, "Transmission of Information," Bell Syst. Tech. J. 7, 535 (1928).

20. R. D. Kell, A. V. Bedford, and M. A. Trainer, "An Experimental Television System," Proc. IRE 22, 1247 (1934); R. D. Kell, A. V. Bedford, and G. L. Fredendall, "A Determination of the Optimum Number of Lines in a Television System," RCA Rev. 5, 8 (1940).

21. D. G. Fink, *Television Engineering Handbook*, (McGraw-Hill Book Co., New York, 1957) p. 8-22.

22. M. W. Baldwin, "The Subjective Sharpness of Simulated Television Images," Bell Syst. Tech. J. 19, 563 (1940).

23. T. N. Cornsweet, *Visual Perception* (Academic Press, New York, 1970), Chapter 11.

24. C. Blakemore and F. W. Campbell, "On the Existence of Neurons in the Human Visual System Selectivity Sensitive to the Orientation and Size of the Retinal Images," J. Physiol. 60, 203 (1969).

25. F. W. Campbell and J. G. Robson, Application of Fourier Analysis to the Visibility of Gratings," J. Physiol. 66, 197 (1968).

26. A. Pantle and R. Sekover, "Size Detecting Mechanisms in Human Vision," Science 162, 1146 (1958).

27. O. H. Schade, "Optical and Photoelectric Analog of the Eye," J. Opt. Soc. Am. 46, 721 (1956).

28. F. L. Van Nes and M. A. Bouman, "Spatial Modulation Transfer in the Human Eye," J. Opt. Soc. Am. 57, 401 (1967).

29. R. Sekuler et al., "Human Visual Motion Sensitivity: Evidence Against a Ratio Theory of Sensory Coding," Percept. Psychophys. 9, 483 (1971).

30. R. W. Cohen (Principal Investigator), "Image Descriptors for Displays-Bimonthly Report Number 5," ONR Contract No. N00014-74-C-0184 (June 1974).

31. D. M. Green and J. A. Swets, *Signal Detection Theory and Psychophysics* (John Wiley and Sons, N.Y., 1966) Chapter 2.